Higher order Moreau's sweeping process: Mathematical formulation and numerical simulation

Vincent Acary, Bernard Brogliato, Daniel Goeleven

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Higher order Moreau's sweeping process: Mathematical formulation and numerical simulation

Vincent Acary — Bernard Brogliato — Daniel Goeleven

N° 5236 – version 2

version initiale June 2004 – version révisée Mai 2005

Thème NUM

Rapport de recherche

N 0249-6399

Higher order Moreau's sweeping process: Mathematical formulation and numerical simulation

Vincent Acary *, Bernard Brogliato † , Daniel Goeleven ‡

Thème NUM — Systèmes numériques
Projet Bipop

Rapport de recherche n° 5236 – version 2§— version initiale June 2004 — version révisée Mai 2005 78 pages

Abstract: In this paper we present an extension of Moreau's sweeping process for higher order systems. The dynamical framework is carefully introduced, and preliminary well-posedness results are given. The time-discretization of these nonsmooth systems with a time-stepping algorithm is also presented. This differential inclusion can be seen as a mathematical formulation of complementarity dynamical systems with arbitrary dimension and arbitrary relative degree between the complementary slackness variables. Applications of such high-order sweeping processes can be found in dynamic optimisation under state constraints and electrical circuits with ideal diodes.

Key-words: Complementarity systems, Hybrid systems, Convex analysis, Measure differential inclusions, Variational inequalities, Numerical simulation, Zero dynamics, Relative degree, Schwartz distributions, Electrical circuits, Time-stepping algorithm, Dissipative systems.

* INRIA Rhône-Alpes, ZIRST Montbonnot, 655 avenue de l'Europe, 38334 Saint-Ismier, France, Vincent.Acary@inrialpes.fr

† INRIA Rhône-Alpes, ZIRST Montbonnot, 655 avenue de l'Europe, 38334 Saint-Ismier, France, Bernard.Brogliato@inrialpes.fr

‡ IREMIA, University of La Réunion, Saint-Denis, 97400, France, Daniel.Goeleven@univ-reunion.f

§ Augmented and revised

Processus de rafle d'ordre supérieur: Formulation mathématique et simulation numérique

Résumé : Cet article propose une extension du processus de rafle de Moreau, au cas de systèmes d'ordre supèrieur à deux. Le cadre dynamique est introduit en détail, et l'existence/unicité des solutions est prouvée. Un schéma numérique de type "time-stepping" est présenté.

Mots-clés : Inéquations variationnelles, inclusions différentielles, systèmes de complémentarité, analyse convexe, distributions de Schwarz, dynamique zéro, degré relatif, algorithme à pas de temps, systèmes dissipatifs.

1 Introduction

The so-called *sweeping process* is a particular differential inclusion of the general form

$$-\dot{z}(t) \in N_{K(t)}(z(t)) \ \ (t \geq 0), \ \ z(0) \in K(0), \tag{1}$$

where $K(t)$ is a convex time-dependent set, and $N_{K(t)}(z(t))$ is the normal cone to $K(t)$ at $z(t)$. Such evolution problems have been introduced by J.J. Moreau in 1971 [48, 49, 50, 51]. These models prove to be quite useful in elastoplasticity, nonsmooth mechanics, convex optimisation, mathematical economics, queuing theory, etc. Generalisations of the sweeping process have been the object of many studies, see e.g. [20, 7, 34, 73, 35, 36, 37, 68, 67, 44] and more references in [45, 19, 38]. Recently it was also shown that quite similar formalisms apply to nonsmooth electrical networks as well as some problems of absolute stability [12, 24].

Moreau (see also Schatzman [63, 64]) introduced [53, 54] an extension of the sweeping process for Lagrangian systems subject to frictionless unilateral constraints. He termed the resulting equations Measure Differential Inclusions. Roughly, such evolution equations are of the form

$$-dv \in N_{V(q(t))}(v(t)), \ \ v(t) \in V(q(t)),$$

where $V(q(t))$ is a closed convex multivalued function depending on $q(\cdot)$, and $v(\cdot)$ is the first derivative of $q(\cdot)$. The term dv denotes the measure that is associated to what one could call the second derivative of $q(\cdot)$. At this stage, we would just like to point out that this evolution problem is of second order, whereas the one in (1) is of first order. This is crucial because it generally implies that the solutions $z(\cdot)$ of the former are absolutely continuous, whereas $v(\cdot)$ in the second formalism is of bounded variation (consequently may possess jumps). In this paper we shall see that the order (which is directly related to some *relative degree* r between two complementary-slackness variables w and λ) indeed is a fundamental parameter which determines the nature of the solutions and their degree as distributions.

Most importantly, numerical time integration schemes have been introduced for both first and second order sweeping processes in [51, 54] which correspond to what is now called *time-stepping* schemes [15]. The term "time-stepping" is used to underline that there is no explicit procedure to take into account events, such as activation of constraints. For the first order case (1), the algorithm is $-z(t_{k+1}) + z(t_k) \in N_{K(t_{k+1})}(z(t_{k+1}))$, and is called the *catching up* algorithm because of its geometrical interpretation. It has been extensively used both for well-posedness studies [45, 38, 43, 70] and for numerical simulation [55, 15]. For Lagrangian systems subject to unilateral constraints, the time integration method is built in the same way and called the Non Smooth Contact Dynamics [54, 56, 32]. For the case $r = 2$, the discretization of the second order sweeping process is suitable [43, 45, 38]. The algorithm that is proposed in this paper, is inspired from these time integration schemes as will be explained in Section 4. One motivation of the proposed study, is to provide us with an efficient numerical scheme for systems with unilateral constraints and with relative degree larger than 3. It is shown in [17] and recalled in the Section 5.5.1 that applying directly a backward Euler scheme leads to some inconsistencies due to the fact that distributional solutions are not correctly approximated. We aim at bridging this gap.

Recently, dynamical complementarity systems (introduced by Moreau in 1963 in the framework of Lagrangian mechanical systems [46, 47], see also [42]) have been the object of interest in the control literature [75, 76, 30, 11, 13, 14, 74, 17, 18, 40, 61, 72], because of many applications in various fields such as nonsmooth electrical networks, optimal control with state and/or input constraint, Lagrangian mechanical systems subject to unilateral constraints, etc [74, 11]. It is well-known that complementarity problems, variational inequalities and inclusions (or generalized equations) are closely related [23]. Similarly, complementarity systems, evolution variational inequalities and differential inclusions as the above ones (which we could name *unbounded DI*) are related [12, 24, 11, 25]. Equivalences between various unilateral dynamics formalisms are established in [1]. For instance it is easy to see that the first order sweeping process can also be rewritten as the evolution variational inequality $\langle \dot{z}(t), v - z(t) \rangle \geq 0$, for all $v \in K(t)$, $z(0) \in K(0)$, $z(t) \in K(t)$ for all $t \geq 0$. Such relationships between various formalisms will be important for the developments in this paper, especially for the design of a numerical algorithm. In parallel results have been obtained for classes of piecewise affine systems [31], however the link with complementarity is not yet clear except in some particular cases [29, §4.2.2]. The time-discretization of linear complementarity systems with implicit Euler algorithms has been considered in [17], as indicated above.

In [59] so-called differential variational inequalities with index ≤ 2 (i.e. with relative degree ≤ 1 in the language of this paper) are studied in a rather detailed manner: well-posedness results are given, and numerical schemes are proposed and shown to converge.

Objective of the paper: The starting point is to consider a dynamical system of the form $\dot{x}(t) = Ax(t) + B\lambda(t)$ whose trajectories have to evolve in a domain of the form $Cx(t) \geq 0$. The dynamics is embedded into a distributional differential inclusion (named the higher order sweeping process) which allows us to integrate the system while respecting the unilateral constraint on the state. In the proposed formalism the Lagrange multiplier λ is a distribution which possesses a specific decomposition into measures which satisfy inclusions into a family of convex cones. A time-stepping numerical algorithm is constructed and its properties are analyzed. Though the convergence towards solutions of the continuous-time system is not yet complete, the analysis shows that the numerical scheme possesses strong properties.

The paper is organised as follows. Some mathematical tools are presented in section 2. A special canonical state space representation which is useful for the subsequent developments is introduced in section 3. The corresponding differential inclusion formalism is introduced and motivated in section 4, where important properties and well-posedness are shown. Section 5 is devoted to the design and the analysis of a time-stepping numerical scheme. In section 6 an application of the extended sweeping process is presented. Conclusions close the paper in section 7.

Notation: The following notation is used: $\mathbb{R}$ is the real line, $\mathbb{R}^+$ is the nonnegative real line, x_i is the ith component of a vector $x \in \mathbb{R}^n$. The relative degree between two signals w and λ defining the output and the input of a system, is denoted as r. The indicator function of a set K is denoted as $\psi_K(\cdot)$, and ∂ is the convex analysis subdifferentiation. The closed convex hull of a set K is denoted by $\overline{\text{co}}(K)$. When K is a nonempty closed convex set then the normal cone to K is denoted as $N_K(\cdot) = \partial\psi_K(\cdot)$. Lexicographical inequalities are denoted as $\succeq$. For a vector x, $x \succeq 0$ means that all entries $x_i = 0$ or the first nonzero entry is > 0. $x \geq 0$ means that all entries $x_i \geq 0$. I_n denotes the $n \times n$ identity matrix, 0^n denotes the vector $(0, 0, \dots, 0)^T \in \mathbb{R}^n$ and $0_n = (0^n)^T$. Let $M \in \mathbb{R}^{n\times n}$ be a symmetric and positive definite matrix, the proximation operator $prox_M[K; .]$ is defined by $prox_M[K; x] = \text{argmin}_{z\in K}(z - x)^T M(z - x)$. If $M = I_n$, we set $prox[K; .] := prox_{I_n}[K; .]$. The notations $\langle x, y\rangle := x^T y$ and $\|x\| := \sqrt{x^T x}$ will also be used.

Acronyms The following acronyms are used throughout the paper: Complementarity Problem (CP), Linear Complementarity Problem (LCP), Variational Inequality (VI), Ordinary Differential Equation (ODE), Differential Algebraic Equation (DAE), Measure Differential Inclusion (MDI), Complementarity System (CS), Linear Complementarity System (LCS).

2 Mathematical Tools

In this section we present fundamental analysis tools which will be helpful in settling the higher order sweeping process formalism. In particular a class of distributions that is a potential class of solutions of the dynamics is presented in detail.

Radon Measure. Let us denote by $\mathcal{B}(\mathbb{R})$ the Borel σ algebra and let dR be a $\mathbb{R}^n-$valued Radon measure, i.e. a Borel regular measure such that $dR_i(K) < +\infty$ for every compact set $K \subset \mathbb{R}$ (see e.g. [22]). Let $A \in \mathcal{B}(\mathbb{R})$ be given, we say that $\{A_i\}_{i=1}^m$ is a finite partition of A if $A_i \in \mathcal{B}(\mathbb{R})$, $A_i \cap A_j = \emptyset, i \neq j$ and $\cup_{i=1}^m A_i = A$. Let us now denote by $\mathcal{P}(A)$ the set of finite partitions of A. The modulus measure of dR is defined by (see e.g. [65]):

$$|dR|(A) = \sup_{\{A_i\}_{i=1}^m \in \mathcal{P}(A)} \sum_{i=1}^m \|dR(A_i)\|, \forall A \in \mathcal{B}(\mathbb{R}).$$

Let $d\mu$ be a real-valued Radon measure. Let us denote by $L^1_{loc}(\mathbb{R}, d\mu; \mathbb{R}^n)$ the space of $d\mu$-locally integrable $\mathbb{R}^n$-valued functions. One says that dR has a density relative to $d\mu$

provided that there exists a (class of) function $R'_\mu \in L^1_{loc}(\mathbb{R}, d\mu; \mathbb{R}^n)$ such that $dR = R'_\mu d\mu$, i.e.

$$dR(A) = \int_A R'_\mu d\mu, \forall A \in \mathcal{B}(\mathbb{R}).$$

If $d\mu$ is nonnegative and dR is absolutely continuous with respect to $d\mu$, i.e. $A \in \mathcal{B}(\mathbb{R}), d\mu(A) = 0 \Rightarrow dR(A) = 0$, then a classical direct consequence of the Lebesgue-Radon-Nikodym theorem (see Theorem 5.10.22 in [65]) ensures the existence of a unique (class of) function $R'_\mu \in L^1_{loc}(\mathbb{R}, d\mu; \mathbb{R}^n)$ such that $dR = R'_\mu d\mu$. Here

$$R'_\mu(t) = \frac{dR}{d\mu}(t),$$

where $\frac{dR}{d\mu}$ denotes the derivative (density) of dR with respect to $d\mu$. In particular, since $|dR|$ is nonnegative and $\|dR(A)\| \leq |dR|(A), \forall A \in \mathcal{B}(\mathbb{R})$, there exists a unique (class of) function $\theta_R \in L^1_{loc}(\mathbb{R}, |dR|; \mathbb{R}^n)$ such that $dR = \theta_R |dR|$.

Differential measure process. Let I be a real non-degenerate interval (not empty nor reduced to a singleton), and let $\{K(t); t \in I\}$ be a family of nonempty closed convex cones. Suppose that $d\mu$ is nonnegative and dR is absolutely continuous with respect to $d\mu$. By convention, we shall write

$$dR = R'_\mu d\mu \in K(t) \text{ on } I \tag{2}$$

to mean that

$$R'_\mu(t) \in K(t), \quad d\mu - a.e. \;\; t \in I. \tag{3}$$

Proposition 1 *If the relation in (3) holds then, for every nonempty bounded $A \in \mathcal{B}(\mathbb{R})$, $A \subset I$:*

$$dR(A) \in \overline{\text{co}}(\cup_{\tau \in A} K(\tau)). \tag{4}$$

Proof: Suppose that (3) is satisfied. If $A \in \mathcal{B}(\mathbb{R})$, $A \subset I$ and $d\mu(A) = 0$ then $dR(A) = 0$ and it is clear that (4) holds since $0 \in \overline{\text{co}}(\cup_{\tau \in A} K(\tau))$. Let $A \in \mathcal{B}(\mathbb{R})$, $A \subset I$ such that $0 < d\mu(A) < +\infty$. Then $R'_\mu(A) \subset \overline{\text{co}}(\cup_{\tau \in A} K(\tau))$ and thus (see Theorem 5.7.35 in [65]):

$$\frac{1}{d\mu(A)} \int_A R'_\mu d\mu \in \overline{\text{co}}(\cup_{\tau \in A} K(\tau))$$

since $\overline{\text{co}}(\cup_{\tau \in A} K(\tau))$ is closed and convex. It results that $dR(A) = \int_A R'_\mu d\mu \in \overline{\text{co}}(\cup_{\tau \in A} K(\tau))$ since $d\mu(A) > 0$ and here $\overline{\text{co}}(\cup_{\tau \in A} K(\tau))$ is a closed convex cone. ■

Functions of bounded variation. Let I denote a non-degenerate real interval (not empty nor reduced to a singleton). By $u \in \mathbf{BV}(\mathbf{I}; \mathbb{R}^\mathbf{n})$ it is meant that u is a $\mathbb{R}^n$-valued function of bounded variation if there exists a constant $C > 0$ such that for all finite sequences $t_0 < t_1 < ... < t_N$ (N arbitrary) of points of I, we have

$$\sum_{i=1}^{N} \|u(t_i) - u(t_{i-1})\| \leq C.$$

Let J be a subinterval of I. The real number $\text{var}(u, J) := \sup \sum_{i=1}^{N} \|u(t_i) - u(t_{i-1})\|$, where the supremum is taken with respect to all the finite sequences $t_0 < t_1 < ... < t_N$ (N arbitrary) of points of J, is called the variation of u in J.

Any BV function has a countable set of discontinuity points and is almost everywhere differentiable. A BV function defined on $[a, b] \subset I$ possesses left-limits in $(a, b]$ and right-limits in $[a, b)$. Moreover, the functions $t \mapsto u(t^+) := \lim_{s \to t, s > t} u(s)$ and $t \mapsto u(t^-) := \lim_{s \to t, s < t} u(s)$ are both BV functions.

We denote by $\mathbf{LBV}(\mathbf{I}; \mathbb{R}^\mathbf{n})$ the space of functions of locally bounded variation, i.e. of bounded variation on every compact subinterval of I.

We denote by $\mathbf{RCLBV}(\mathbf{I}; \mathbb{R}^\mathbf{n})$ the space of right-continuous functions of locally bounded variation. It is known that if $u \in RCLBV(I; \mathbb{R}^n)$ and $[a, b]$ denotes a compact subinterval of I, then $u(t)$ can be represented in the form (see e.g. [66]):

$$u(t) = \mathcal{J}_u(t) + [u](t) + \zeta_u(t), \forall t \in [a, b],$$

where $\mathcal{J}_u$ is a jump function, $[u]$ is an absolutely continuous function and ζ_u is a singular function. Here $\mathcal{J}_u$ is a jump function in the sense that $\mathcal{J}_u$ is right-continuous and given any $\varepsilon > 0$, there exists finitely many points of discontinuity $t_1, ..., t_N$ of $\mathcal{J}_u$ such that $\sum_{i=1}^{N} \|\mathcal{J}_u(t_i) - \mathcal{J}_u(t_i^-)\| + \varepsilon > \text{var}(\mathcal{J}_u, [a, b])$, $[u]$ is an absolutely continuous function in the sense that for every $\varepsilon > 0$, there exists $\delta > 0$ such that $\sum_{i=1}^{N} \|[u](\beta_i) - [u](\alpha_i)\| < \varepsilon$, for any collection of disjoint subintervals $(\alpha_i, \beta_i] \subset [a, b] (1 \leq i \leq N)$ such that $\sum_{i=1}^{N} (\beta_i - \alpha_i) < \delta$, and ζ_u is a singular function in the sense that ζ_u is a continuous and bounded variation function on $[a, b]$ such that $\dot{\zeta}_u = 0$ almost everywhere on $[a, b]$.

By $u \in \mathbf{RCSLBV}(\mathbf{I}; \mathbb{R}^\mathbf{n})$ it is meant that u is a right-continuous function of special locally bounded variation, i.e. u is of bounded variation and can be written as the sum of a jump function and an absolutely continuous function on every compact subinterval of I. So, if $u \in \mathbf{RCSLBV}(\mathbf{I}; \mathbb{R}^\mathbf{n})$ then

$$u = [u] + \mathcal{J}_u \tag{5}$$

where $[u]$ is a locally absolutely continuous function called the absolutely continuous component of u and $\mathcal{J}_u$ is uniquely defined up to a constant by

$$\mathcal{J}_u(t) = \sum_{t \geq t_n} u(t_n^+) - u(t_n^-) = \sum_{t \geq t_n} u(t_n) - u(t_n^-) \tag{6}$$

where $t_1 < t_2 < ... < t_n < ...$ denote the countably many points of discontinuity of u in I. given in (6).

Stieltjes measure. Let $u \in LBV(I;\mathbb{R}^n)$ be given. We denote by du the Stieltjes measure generated by u (see e.g. [65] and [45]). Recall that for $a \leq b$, $a, b \in I$:

$$du([a,b]) = u(b^+) - u(a^-),$$

$$du([a,b)) = u(b^-) - u(a^-),$$
$$du((a,b]) = u(b^+) - u(a^+),$$
$$du((a,b)) = u(b^-) - u(a^+).$$

In particular, we have

$$du(\{a\}) = u(a^+) - u(a^-).$$

For $u \in LBV(I;\mathbb{R}^n)$, u^+ and u^- denote the functions defined by

$$u^+(t) = u(t^+) = \lim_{s \to t, s > t} u(s)), \forall t \in I,$$

and

$$u^-(t) = u(t^-) = \lim_{s \to t, s < t} u(s)), \forall t \in I.$$

If $u, v \in LBV(I;\mathbb{R}^n)$ then $u^T v \in LBV(I;\mathbb{R})$ and

$$d(u^T v) = (v^-)^T du + (u^+)^T dv = (v^+)^T du + (u^-)^T dv. \tag{7}$$

Let us also recall that

$$2(u^-)^T du \leq d(u^T u) = (u^+ + u^-)^T du \leq 2(u^+)^T du. \tag{8}$$

Measure differential inclusions. The material given in this Section can be used to formulate measure differential inclusions. For example, let $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ be a C^1-mapping and let $C \subset \mathbb{R}^n$ be a nonempty closed convex set. We consider the measure differential inclusion: Find $u \in RCLBV(\mathbb{R};\mathbb{R}^n)$ such that:

$$du + F(t, u(t))dt \in -N_C(u(t)). \tag{9}$$

The sense of (9) is given by the existence of a nonnegative Radon measure $d\mu$ such that the measures du and dt are absolutely continuous with respect to $d\mu$ and

$$\frac{du}{d\mu}(t) + F(t, u(t))\frac{dt}{d\mu}(t) \in -N_C(u(t)), \quad d\mu - \text{a.e. } t \in \mathbb{R}. \tag{10}$$

Note that the concept of solution does not depend of the choice of the nonnegative Radon measure $d\mu$ since the right-hand side of (10) is a cone, and the densities can be obtained one from other by multiplication with a nonnegative function. It results from Proposition 1 that

$$du(A) + \int_A F(\tau, u(\tau))d\tau \in \overline{\text{co}}(\bigcup_{\tau \in A} -N_C(u(\tau))), \tag{11}$$

for every $A \in \mathcal{B}(\mathbb{R})$ such that $d\mu(A) < +\infty$.

Let us now recall some important consequences of such a mathematical model. We have

$$u(t^+) - u(t^-) \in -N_C(u(t)), \forall t \in \mathbb{R} \tag{12}$$

and

$$u_t' + F(t, u(t)) \in -N_C(u(t)), \ \ a.e. \ \ t \in \mathbb{R}.$$

According to the notations introduced above, u_t' denotes the density of du with respect to the Lebesgue measure dt.

Distributions generated by RCSLBV functions. Let I be the open real interval given by

$$I =]\alpha, \beta[$$

where $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R} \cup \{+\infty\}$. We set also

$$\tilde{I} = [\alpha, \beta[.$$

Here $C_0^\infty(I)$ denotes the space of real-valued $C^\infty(I)$-mappings with compact support and $\mathcal{D}'(I)$ is the space of Schwartz distributions on I. Recall that for $T \in \mathcal{D}'(I)$, the (generalized) derivative of T is defined by

$$\langle DT, \varphi \rangle = -\langle T, \dot{\varphi} \rangle, \forall \ \varphi \in C_0^\infty(I).$$

The (generalized) derivative of order n is then given by

$$D^n T = D(D^{n-1}T) \ (n \geq 2),$$

that is

$$\langle D^n T, \varphi \rangle = (-1)^n \langle T, \varphi^{(n)} \rangle, \forall \ \varphi \in C_0^\infty(I).$$

For $a \in I$, we denote by δ_a the Dirac distribution at a, defined by

$$\langle \delta_a, \varphi \rangle = \varphi(a), \forall \ \varphi \in C_0^\infty(I).$$

Note that $\delta_a = D\mathcal{H}(. - a)$ where $\mathcal{H}$ is the Heaviside function:

$$\mathcal{H}(t) = \begin{cases} 1 & \text{if} \ \ t \geq 0 \\ 0 & \text{if} \ \ t < 0 \end{cases}.$$

The support $\mathrm{supp}\{\varphi\}$ of a function $\varphi : I \rightarrow \mathbb{R}$ is defined by $\mathrm{supp}\{\varphi\} := \overline{\{t \in I : \varphi(t) \neq 0\}}$. The support $\mathrm{supp}\{T\}$ of a distribution $T \in \mathcal{D}'(I)$ is defined by $\mathrm{supp}\{T\} := I \backslash \mathcal{O}$ where $\mathcal{O} \subset I$ denotes the largest open set of I on which T vanishes in the sense that $\langle T, \bar{\varphi} \rangle = 0, \forall\ \varphi \in C_0^\infty(\mathcal{O})$. Here, for $\varphi \in C_0^\infty(\mathcal{O})$, the function $\bar{\varphi} \in C_0^\infty(I)$ is defined by $\bar{\varphi}(t) = \begin{cases} \varphi(t) & \text{if } t \in \mathcal{O} \\ 0 & \text{if } t \in I \backslash \mathcal{O} \end{cases}$.

Let $h \in RCSLBV(\tilde{I};\mathbb{R})$ be given. We will denote by $E_0(h)$ the countable set of points of discontinuity $t_1 < t_2 < ... < t_k < ...$ of h. As seen above, h can be written as the sum of a locally absolutely continuous function $[h]$ and the locally jump function J_h given by

$$J_h(t) = \sum_{t \geq t_k} \sigma_h(t),$$

where for $t \in \tilde{I}$,

$$\sigma_h(t) := h(t^+) - h(t^-) = h(t) - h(t^-)$$

denotes the jump of h at t. It is clear that if $t \in \tilde{I} \backslash E_0(h)$ then $\sigma_h(t) = 0$.

We will denote by $\hat{h}^{(1)}$ the right derivative of the absolutely continuous part $[h]$ of $h \in RCSLBV(\tilde{I};\mathbb{R})$, i.e.

$$\hat{h}^{(1)}(t) := \frac{d^+[h]}{dt}(t) = \lim_{\sigma \to 0^+} \frac{[h](t+\sigma) - [h](t)}{\sigma}$$

We have thus:

$$h = [h] + J_h \tag{13}$$

and

$$dh = \hat{h}^{(1)} dt + d\mathcal{J}_h. \tag{14}$$

The measure $d\mathcal{J}_h$ is atomic as a measure concentrated on the set E_0 of countably many points of discontinuity of h in $\tilde{I}$, i.e. $d\mathcal{J}_h(A) = 0, \forall A \in \mathcal{B}(\mathbb{R}), A \subset \mathbb{R} \backslash E_0$.

Let us now set

$$\mathcal{F}_0(\tilde{I};\mathbb{R}) = RCSLBV(\tilde{I};\mathbb{R})$$

$$\mathcal{F}_1(\tilde{I};\mathbb{R}) = \{h \in \mathcal{F}_0(\tilde{I};\mathbb{R}) : \hat{h}^{(1)} \in RCSLBV(\tilde{I};\mathbb{R})\}$$

$$\mathcal{F}_2(\tilde{I};\mathbb{R}) = \{h \in \mathcal{F}_1(\tilde{I};\mathbb{R}) : \hat{h}^{(2)} := \frac{d^+}{dt}[\hat{h}^{(1)}] \in RCSLBV(\tilde{I};\mathbb{R})\}$$

$$.....$$

$$\mathcal{F}_k(\tilde{I};\mathbb{R}) = \{h \in \mathcal{F}_{k-1}(\tilde{I};\mathbb{R}) : \hat{h}^{(k)} := \frac{d^+}{dt}[\hat{h}^{(k-1)}] \in RCSLBV(\tilde{I};\mathbb{R})\}$$

and

$$\mathcal{F}_\infty(\tilde{I};\mathbb{R}) = \cap_{k \in \mathbb{N}} \mathcal{F}_k(\tilde{I};\mathbb{R}).$$

We standardize the notations by setting $\hat{h}^{(0)} := h$. Note that $\hat{h}^{(\alpha)} \in RCSLBV(\tilde{I};\mathbb{R})$ $(\alpha \geq 1)$ means that the absolutely continuous function $[\hat{h}^{(\alpha-1)}]$ admits a right derivative $\hat{h}^{(\alpha)} = \frac{d^+}{dt}[\hat{h}^{(\alpha-1)}]$ at each $t \in \tilde{I}$ and $\hat{h}^{(\alpha)}$ has special local bounded variation over $\tilde{I}$.

Remark 1 *i) Note that if* $\frac{d^+}{dt}[\hat{h}^{(\alpha-1)}]$ *exists and has special local bounded variation on* $\tilde{I}$ *then necessarily* $\frac{d^+}{dt}[\hat{h}^{(\alpha-1)}]$ *is right-continuous. Properties* $\hat{h}^{(\alpha)} \in RCLBV(\tilde{I};\mathbb{R})$ *and* $\hat{h}^{(\alpha)} \in LBV(\tilde{I};\mathbb{R})$ *are thus equivalent. The "RC" requirement stands in the definition of the vector space* $\mathcal{F}_\alpha(\tilde{I};\mathbb{R})$ *for pedagogical reasons.*

ii) Note also that for a.e. $t \in \tilde{I}$, *we have* $\frac{d^+}{dt}[\hat{h}^{(\alpha-1)}](t) = \frac{d^+}{dt}\hat{h}^{(\alpha-1)}(t) = h_t^{(\alpha-1)}(t)$

Example 1 *Set* $I =]0,+\infty[$ *and let* $u : \tilde{I} \to \mathbb{R}$ *be the function given by*

$$u(t) = |\sin(t)|, \quad \forall t \geq 0.$$

It is clear that

$$\hat{u}^{(0)} := u \in RCSLBV(\tilde{I};\mathbb{R})$$

since u *is Lipschitz-continuous. Then we obtain*

$$\hat{u}^{(1)}(t) := \frac{d^+}{dt}[\hat{u}^{(0)}](t) = \frac{d^+u}{dt}(t) = \cos(t-k\pi) \ \text{ if } \ t \in [k\pi,(k+1)\pi), \ (k \in \mathbb{N}).$$

We see that $E_0(\hat{u}^{(1)}) = \{k\pi;\ k \in \mathbb{N}\backslash\{0\}\}$ *and*

$$\hat{u}^{(1)} = [\hat{u}^{(1)}] + J,$$

where

$$[\hat{u}^{(1)}](t) = -2k + cos(t-k\pi) \ \text{ if } \ t \in [k\pi,(k+1)\pi], \ (k \in \mathbb{N})$$

and

$$J(t) = 2k \ \text{ if } \ t \in [k\pi,(k+1)\pi], \ (k \in \mathbb{N}).$$

Thus

$$\hat{u}^{(1)} \in RCSLBV(\tilde{I};\mathbb{R})$$

Then

$$\hat{u}^{(2)}(t) := \frac{d^+}{dt}[\hat{u}^{(1)}](t) = -|\sin(t)|$$

so that

$$\hat{u}^{(2)} \in RCSLBV(\tilde{I};\mathbb{R}).$$

And so on, we see that

$$\hat{u}^{(k)}(t) = \begin{cases} (-1)^m\hat{u}^{(0)}(t) & \text{if } \ k = 2m \\ (-1)^m\hat{u}^{(1)}(t) & \text{if } \ k = 2m+1 \end{cases}, \quad (m \in \mathbb{N}),$$

so that $\hat{u}^{(k)} \in RCSLBV(\tilde{I};\mathbb{R}),\ \forall k \in \mathbb{N}$, *and thus*

$$u \in \mathcal{F}_\infty(\tilde{I};\mathbb{R}).$$

Let $h \in \mathcal{F}_\infty(\tilde{I}; \mathbb{R})$ be given. One remarks that (generalized) derivatives of h have a specific easy to handle structure. Indeed, let us here denote by T_h the regular distribution generated by h, i.e.

$$\langle T_h, \varphi \rangle = \int_I \varphi \; h \; dt, \; \forall \; \varphi \in C_0^\infty(I).$$

Let $\varphi \in C_0^\infty(I)$ be given. Using (7) and (14), we obtain (here $h^+ = h$ and $\varphi^- = \varphi$):

$$\langle DT_h, \varphi \rangle = -\int_I \dot{\varphi} \; h \; dt = -\int_I \; h \; d\varphi = -\int_I d(h\varphi) + \int_I \varphi \; dh$$

$$= \int_I \varphi \; dh = \int_I \varphi \; \hat{h}^{(1)} dt + \sum_{t_k \in E_0(h) \cap \mathrm{supp}\{\varphi\}} (h(t_k^+) - h(t_k^-)) \langle \delta_{t_k}, \varphi \rangle$$

and so on, we see that for any $\alpha \geq 2$:

$$\langle D^\alpha T_h, \varphi \rangle = \int_I \varphi \; d\hat{h}^{(\alpha-1)} + \sum_{i=2}^{\alpha} \Big(\sum_{t_k \in E_0(\hat{h}^{(\alpha-i)}) \cap \mathrm{supp}\{\varphi\}} (\hat{h}^{(\alpha-i)}(t_k^+) - \hat{h}^{(\alpha-i)}(t_k^-)) \langle \delta_{t_k}^{(i-1)}, \varphi \rangle \Big)$$

$$= \int_I \varphi \; \hat{h}^{(\alpha)} dt + \sum_{i=1}^{\alpha} \Big(\sum_{t_k \in E_0(\hat{h}^{(\alpha-i)}) \cap \mathrm{supp}\{\varphi\}} (\hat{h}^{(\alpha-i)}(t_k^+) - \hat{h}^{(\alpha-i)}(t_k^-)) \langle \delta_{t_k}^{(i-1)}, \varphi \rangle \Big).$$

Example 2 *Let $u : \tilde{I} \to \mathbb{R}$ be the function considered in Example 1. Let us now consider the distribution T defined by*

$$\langle T, \varphi \rangle = \int_{-\infty}^{\infty} |\sin(t)| \varphi(t) dt, \; \forall \; \varphi \in C_0^\infty(I).$$

Then for a given function $\varphi \in C_0^\infty(I)$, we see that:

$$\langle DT, \varphi \rangle = \int_{-\infty}^{\infty} \hat{u}^{(1)}(t) \varphi(t) dt = \sum_{k \in \mathbb{N} \cap \mathrm{supp}\{\varphi\}} \int_{k\pi}^{(k+1)\pi} \cos(t - k\pi) \varphi(t) dt,$$

$$\langle D^2 T, \varphi \rangle = \int_{-\infty}^{\infty} \hat{u}^{(2)}(t) \varphi(t) dt + \sum_{k \in \mathbb{N} \setminus \{0\} \cap \mathrm{supp}\{\varphi\}} (\hat{u}^{(1)}(k\pi^+) - \hat{u}^{(1)}(k\pi^-)) \langle \delta_{k\pi}, \varphi \rangle =$$

$$= -\int_{-\infty}^{\infty} |\sin(t)| \varphi(t) dt + 2 \sum_{k \in \mathbb{N} \setminus \{0\} \cap \mathrm{supp}\{\varphi\}} \langle \delta_{k\pi}, \varphi \rangle,$$

and so on.

Let us now denote by $\mathcal{T}_n(I)$ the set of all distributions $T \in \mathcal{D}'(I)$ which satisfy the following condition: there exists a function $F \in \mathcal{F}_\infty(\tilde{I};\mathbb{R})$ such that $T = D^n F$. It is clear that $\mathcal{T}_0(I) = \mathcal{F}_\infty(\tilde{I};\mathbb{R})$. If $T \in \mathcal{T}_1(I)$ then there exists $F \in \mathcal{F}_\infty(\tilde{I};\mathbb{R})$ such that $\langle T, \varphi\rangle = \int_I \varphi d\hat{F}^{(0)} = \int_I \varphi dF$, for all $\forall \varphi \in C_0^\infty(I))$. More generally, if $T \in \mathcal{T}_n(I)$ for some $n \geq 2$, then there exists $F \in \mathcal{F}_\infty(I;\mathbb{R})$ such that, for all $\varphi \in C_0^\infty(I)$:

$$\langle T,\varphi\rangle = \int_I \varphi \, d\hat{F}^{(n-1)} + \sum_{i=2}^{n} \Big(\sum_{t_k \in E_0(\hat{F}^{(n-i)}) \cap \mathrm{supp}\{\varphi\}} (\hat{F}^{(n-i)}(t_k^+) - \hat{F}^{(n-i)}(t_k^-)) \langle \delta_{t_k}^{(i-1)}, \varphi\rangle \Big)$$

$$= \int_I \varphi \, \hat{F}^{(n)} dt + \sum_{i=1}^{n} \Big(\sum_{t_k \in E_0(\hat{F}^{(n-i)}) \cap \mathrm{supp}\{\varphi\}} \sigma_{\hat{F}^{(n-i)}}(t_k) \langle \delta_{t_k}^{(i-1)}, \varphi\rangle \Big).$$

For such distribution T, we may clearly identify the "function part" $\{T\}$ and the "measure part" $\ll T \gg$ respectively by

$$\{T\} = \hat{F}^{(n)}$$

and (if $n \geq 1$)

$$\langle \ll T \gg, \varphi\rangle = \int_I \varphi \, d\hat{F}^{(n-1)}, \ \forall \, \varphi \in C_0^\infty(I).$$

We will also use the notation $d \ll \mathrm{T} \gg$ to denote the Stieltjes measure $d\hat{F}^{(n-1)}$ generated by $\hat{F}^{(n-1)} \in RCSLBV(\tilde{I};\mathbb{R})$. Here $\{T\}$ is a RCSLBV function and $d \ll \mathrm{T} \gg$ is a Stieltjes measure. For pedagogical reasons, we use the two different notations $d \ll \mathrm{T} \gg$ and $\ll T \gg$ to denote respectively the Radon measure defined on the Borel's sets and the corresponding distribution, i.e.

$$\langle \ll T \gg, \varphi\rangle = \int_I \varphi \; d \ll \mathrm{T} \gg, \ \ \forall \, \varphi \in C_0^\infty(I).$$

It will be also convenient to use the notation $\{T^{(k)}\}$ to denote the "function part" of $D^k T$, i.e.

$$\{T^{(k)}\} = \{D^k T\} = \hat{F}^{(n+k)}.$$

For $T \in \cup_{n \in \mathbb{N}} \mathcal{T}_n(I)$, we define the degree "deg(T)" of T in the following way: Let n be the smallest integer such that $T \in \mathcal{T}_n(I)$, we set

$$\deg(T) = \begin{cases} n+1 \ \text{ if } \ n \geq 1 \\ 1 \ \text{ if } \ n = 0 \ \text{ and } \ E_0(\{T\}) \neq \emptyset \\ 0 \ \text{ if } \ n = 0 \ \text{ and } \ E_0(\{T\}) = \emptyset \end{cases} \tag{15}$$

The distributions of degree 0 are the continuous $\mathcal{F}_\infty$-functions while the distributions of degree 1 are the discontinuous $\mathcal{F}_\infty$-functions. The right-continuous Heaviside function is of degree 1, the Dirac distribution δ is of degree 2, the distribution $\delta^{(n)}$ is of degree $n+1$.

Example 3 *Let us here consider the distribution T of Example 2. We have:*

$$T \equiv \{T\} = \hat{u}^{(0)} = |\sin(.)|, \quad \deg(T) = 0,$$

$$DT \equiv \{T^{(1)}\} = \{DT\} = \hat{u}^{(1)} = \cos(. - k\pi) \text{ on } [k\pi, (k+1)\pi) \ (k \in \mathbb{N}), \quad \deg(DT) = 1,$$

$$D^2T \equiv \ll D^2T \gg = -|\sin(.)| + 2 \sum_{k \in \mathbb{N}\setminus\{0\}} \delta_{k\pi}, \quad \deg(D^2T) = 2,$$

$$\{T^{(2)}\} = \{D^2T\} = \hat{u}^{(2)} = -|\sin(.)|,$$

and

$$d \ll D^2T \gg = d\hat{u}^{(1)}.$$

3 The ZD canonical representation

In this section a canonical state space representations is derived, which will prove to be useful to formalize the extended sweeping process. Our treatment here is, of necessity informal. The scope of this section is not to get all the smoothness hypotheses worked out but is merely to illustrate a canonical state representation that is generally used by researchers from Systems and Control.

3.1 Canonical state space representation

Let us here consider the following dynamical system:

$$\begin{cases} \dot{x}(t) = Ax(t) + B\lambda(t) \ \ (t \geq 0) \\ \\ x(0) = x_0 \\ \\ w(t) = Cx(t) \geq 0 \ \ (t \geq 0) \end{cases} \tag{16}$$

where $x : \mathbb{R}^+ \to \mathbb{R}^n$, $\lambda \in \mathbb{R}$, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$ and $C \in \mathbb{R}^{1\times n}$. We assume that the transfer function $H(s) = C(sI_n - A)^{-1}B$ of the operator $\lambda \mapsto w$ is not zero. Equivalently there exists $1 \leq r \leq n$ such that $CA^{r-1}B \neq 0$ while $CA^{i-1}B = 0$ for all $1 \leq i \leq r-1$. The integer r is called the relative degree of the triple (A, B, C). Let us note that it is implicitly assumed in (16) that the relative degree is strictly larger than zero. Actually, the framework that is presented next is essentially linked to systems with $r \geq 1$. The existence of a relative degree allows one to perform a state space transformation, with new state vector $z = Wx$, W square full-rank of order n, and

$$z^T = (w, \dot{w}, \ddot{w}, ..., w^{(r-1)}, \xi^T) = (\bar{z}^T, \xi^T), \ \xi \in \mathbb{R}^{n-r} \tag{17}$$

such that the new state space representation is (see [77]):

$$\left\{\begin{array}{l}\dot{z}_1(t)=z_2(t)\ \ (t\geq 0)\\ \dot{z}_2(t)=z_3(t)\ \ (t\geq 0)\\ \dot{z}_3(t)=z_4(t)\ \ (t\geq 0)\\ \\ \vdots\\ \\ \dot{z}_{r-1}(t)=z_r(t)\ \ (t\geq 0)\\ \dot{z}_r(t)=CA^rW^{-1}z(t)+CA^{r-1}B\lambda(t)\ \ (t\geq 0)\\ \\ \dot{\xi}(t)=A_\xi\xi(t)+B_\xi z_1(t)\ \ (t\geq 0)\\ \\ w(t)=z_1(t)\ \ (t\geq 0)\\ \\ z(0)=z_0.\end{array}\right. \tag{18}$$

Here $A_\xi \in \mathbb{R}^{n-r\times n-r}$ and $B_\xi \in \mathbb{R}^{n-r\times 1}$. The transition matrix of the ZD form is

$$WAW^{-1}=\begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0_{n-r}\\ 0 & 0 & 1 & \dots & 0 & 0_{n-r}\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ 0 & 0 & \dots & \dots & 1 & 0_{n-r}\\ d_1 & d_2 & d_3 & \dots & d_r & d_\xi^T\\ B_\xi & 0^{n-r} & 0^{n-r} & \dots & 0^{n-r} & A_\xi \end{pmatrix} \tag{19}$$

where $(0^{n-r})^T=0_{n-r}=(0,\dots,0)\in\mathbb{R}^{1\times(n-r)}$. Moreover

$$CA^rW^{-1}z=d^T\bar{z}+d_\xi^T\xi, \tag{20}$$

with $d^T=(d_1,\dots,d_r)$.

In Systems and Control theory, the dynamics $\dot{\xi}=A_\xi\xi$ is called the *zero dynamics*, so we shall denote the state space form in (18) the ZD representation. We note that the formalism in (18) continues to hold if $\lambda, w\in\mathbb{R}^m$. The system has a uniform relative degree if all the Markov parameter $C_iA^{r-1}B_j$ are non zero for some r and all integers $i,j\in\{1,...,m\}$, where

C_i is the ith row of C while B_j is the jth column of B. Then $CA^{r-1}B$ is an $m \times m$ matrix. One has in the multi-variable case $z_1^i(\cdot) = w_i(\cdot)$, $1 \leq i \leq m$ and

$$\left\{\begin{array}{l} \dot{z}_1^i(t) = z_2^i(t) \quad (t \geq 0) \\ \dot{z}_2^i(t) = z_3^i(t) \quad (t \geq 0) \\ \vdots \\ \dot{z}_r^i(t) = C_i A^r W^{-1} z(t) + C_i A^{r-1} B\lambda(t) \quad (t \geq 0) \\ \\ \dot{\xi}^i(t) = A_{\xi^i}\xi^i(t) + B_{\xi^i} z_1^i(t) \quad (t \geq 0) \\ \\ w(t) = z_1(t) \quad (t \geq 0). \end{array}\right. \tag{21}$$

Grouping terms together $z_1 = (z_1^1, \ldots, z_1^m)^T$, and so on, one gets the same expression as in (18) but all z_i, $1 \leq i \leq r$, are $m-$dimensional. One has also $mr \leq n$.

Example 4 *Let the linear part of a system as in (16) be described by its transfer function* $H(s) = \frac{w(s)}{\lambda(s)} = \frac{s^2-1}{s^4+s^3-(1+\alpha)s-1-\beta}$. *The ZD canonical state space representation is*

$$\left\{\begin{array}{l} \dot{z}_1(t) = z_2(t) \quad (t \geq 0) \\ \dot{z}_2(t) = \lambda(t) - z_1(t) - z_2(t) + \alpha\xi_1(t) + \beta\xi_2(t) \quad (t \geq 0) \\ \dot{\xi}_1(t) = \xi_2(t) \quad (t \geq 0) \\ \dot{\xi}_2(t) = \xi_1(t) + z_1(t) \quad (t \geq 0) \\ w(t) = z_1(t) \quad (t \geq 0) \end{array}\right. \tag{22}$$

3.2 Distributional dynamics model

Starting from this section on, we set $I =]0, +\infty[$, and only functions from the class $RCSLBV(\tilde{I}; \mathbb{R})$ and distributions from the class $\cup_{n\in\mathbb{N}}\mathcal{T}_n(I)$ are used. Moreover, in order to simplify the presentation of our problem, we shall assume that $m = 1$. When the statements or results also obviously hold for $m \geq 2$ and uniform relative degree (see (21)) this will be pointed out.

In this section, we present our model. Recalling that all concrete mathematical models were born without mathematical maturity, we confess that it is also the case of the following one. We need thus in this section to stay beyond the standards required in the mathematical literature. Happily, this will be brief, and our aim in the following sections will be precisely to bring our model to come to its mathematical maturity.

Until now possible state $x(\cdot)$ jumps have not been introduced. It is of utmost importance to notice that in general, the solutions of (18) (equivalently of (16)) will not be differentiable. Consider for instance the initial data $z_i(0^-) \leq -\delta$ for some $\delta > 0$ and all $1 \leq i \leq r$. Then obviously all the z_i, $1 \leq i \leq r$, need to jump to some non-negative value so that the unilateral constraint $z_1(t) \geq 0$ is satisfied on $(0, \epsilon)$ for some $\epsilon > 0$. At this stage we can just say that a jump mapping is needed. Its form will depend on the type of system one handles (in Mechanics, this is the realm of impact mechanics [10]). If one considers (18) as an equality of distributions of class $\cup_{n\in\mathbb{N}}\mathcal{T}_n(I)$, then we can rewrite it as

$$\left\{\begin{array}{l} Dz_1 = z_2 \\ Dz_2 = z_3 \\ Dz_3 = z_4 \\ \\ \vdots \\ \\ Dz_{r-1} = z_r \\ Dz_r = CA^rW^{-1}z + CA^{r-1}B\lambda \\ \\ D\xi = A_\xi \xi + B_\xi z_1. \end{array}\right. \tag{23}$$

At a reinitialisation time one has $\{z\}(t_k^+) = \mathcal{F}(\{z\}(t_k^-))$, where $\mathcal{F}(\cdot)$ is an operator that will be defined later. Consider the above initial conditions on $\{z\}(\cdot)$. Then Dz_1 is a distribution of degree 2 and we get $Dz_1 = \{\dot{z}_1\} + \sigma_{z_1}(0)\delta_0 = z_2$. Consequently Dz_2 is a distribution of degree 3 and $Dz_2 = D^2z_1 = D\{\dot{z}_1\} + \sigma_{z_1}(0)\dot{\delta}_0 = \{\dot{z}_2\} + \sigma_{\{\dot{z}_1\}}(0)\delta_0 + \sigma_{z_1}(0)\dot{\delta}_0 = z_3$, and $\{\dot{z}_1\} = \{z_2\}$. Then Dz_3 is a distribution of degree 4, and we get $Dz_3 = D\{\dot{z}_2\} + \sigma_{\{\dot{z}_1\}}(0)\dot{\delta}_0 + \sigma_{z_1}(0)\ddot{\delta}_0 = \{\dot{z}_3\} + \sigma_{\{\dot{z}_2\}}(0)\delta_0 + \sigma_{\{\dot{z}_1\}}(0)\dot{\delta}_0 + \sigma_{z_1}(0)\ddot{\delta}_0 = z_4$, and $\{\dot{z}_2\} = \{z_3\}$. Thus $\sigma_{\{\dot{z}_1\}}(0) = \{z_2\}(0^+) - \{z_2\}(0^-)$, and $\sigma_{\{\dot{z}_2\}}(0) = \{z_3\}(0^+) - \{z_3\}(0^-)$. And so on. Until now we have decomposed only the left hand side of the dynamics as distributions of some degrees. Now let us get back to the distributional dynamics in (23). Starting from $Dz_1 = z_2$, one deduces that the right hand side has to be of the same degree than the left hand side. This means that the right hand side is equal to $\{z_2\} + \nu_1$, where ν_1 is a distribution of degree 2, i.e. a measure. Similarly from $Dz_2 = z_3$ one deduces that $z_3 = \{z_3\} + \nu_2'$, where ν_2' has degree 3 and can therefore further be decomposed as $\nu_2 + \nu_1'$, with $\deg(\nu_2) = 2$ and $\deg(\nu_1') = 3$. It is not difficult to see that $\nu_1' = D\nu_1$. Therefore $Dz_2 = \{z_3\} + \nu_2 + D\nu_1$. The variables ν_1 and ν_2 are slack variables (or Lagrange multipliers), and are measures of the form $\nu_i(\varphi) = \int_I \varphi(t)d\nu_i \;\; \forall\varphi \in C_0^\infty(I)$, where $d\nu_i$ is a Stieltjes measure generated by a $\mathcal{F}_\infty(\tilde{I};\mathbb{R})$-

function. Continuing the reasoning until Dz_r, we obtain $Dz_r = CA^rW^{-1}\{z\} + CA^{r-1}B\lambda$ where deg(λ) =deg(Dz_r) $= r+1$. Consequently from (23) one gets

$$\left\{\begin{array}{l} Dz_1 = \{z_2\} + \nu_1 \\ Dz_2 = \{z_3\} + D\nu_1 + \nu_2 \\ Dz_3 = \{z_4\} + D^2\nu_1 + D\nu_2 + \nu_3 \\ \\ \vdots \\ Dz_i = \{z_{i+1}\} + D^{(i-1)}\nu_1 + D^{(i-2)}\nu_2 + \ldots + D\nu_{i-1} + \nu_i \\ \vdots \\ Dz_{r-1} = \{z_r\} + D^{(r-2)}\nu_1 + \ldots + \nu_{r-1} \\ Dz_r = CA^rW^{-1}\{z\} + CA^{r-1}B\lambda. \end{array}\right. \tag{24}$$

We keep the notation λ for the multiplier which appears in the last line. One sees that λ in (24) can be given a meaning as

$$\lambda = (CA^{r-1}B)^{-1}[D^{(r-1)}\nu_1 + \ldots + D\nu_{r-1}] + \nu_r \tag{25}$$

provided $CA^{r-1}B \neq 0$ (invertible in the case $m \geq 2$). Then λ is uniquely defined as in (25).

It is important at this stage to realise that λ is the unique source of higher degree distributions in the system, which will allow the state to jump. Therefore the measures ν_i have themselves to be considered as sub-multipliers. In (24) we have separated the regular (functions) parts denoted as $\{\cdot\}$ (see Section 2) and the atomic distributional parts. Notice that $\{z_{i+1}\} = \{\dot{z}_i\}$. Also $D\{z_i\} = \{z_{i+1}\} + \nu_i$ (see Section 2).

Only the Dirac measures ν_i and time functions are signed. Consequently imposing $\lambda \geq 0$ is meaningless in general. Another point of view is to assert that $\lambda \geq 0$ implies that λ is a measure. However this is not sufficient to assure $z_1 \geq 0$ along the time integration. Consequently one has to resort to higher degree distributions to give a reasonably general meaning to the dynamics in (16).

Our aim is now to propose a mathematical formulation applicable to the study of our problem. Let us here also suppose that $CA^{r-1}B \neq 0$ (invertible if $m \geq 2$) .

Distributional formalism. A mathematical problem that appears suitable for the study of the dynamics in (23), (24), (25) consists to find $z_1, ..., z_r \in \mathcal{T}_{r-1}(\mathbb{R}^+)$ and $\xi_i \in \mathcal{T}_{r-1}(\mathbb{R}^+)$ $(1 \leq$

$i \leq n-r$) satisfying the distributional equations

$$\left\{\begin{array}{l} Dz_1 - z_2 = 0 \\ Dz_2 - z_3 = 0 \\ Dz_3 - z_4 = 0 \\ \\ \vdots \\ \\ Dz_{r-1} - z_r = 0 \\ Dz_r - CA^r W^{-1} z = CA^{r-1} B\lambda \\ \\ D\xi = A_\xi \xi + B_\xi z_1 \end{array}\right. \tag{26}$$

together with some constraints on the distributions (of degree ≤ 2) $\nu_1, ..., \nu_r$ defined by

$$\left\{\begin{array}{l} \nu_1 := \ll Dz_1 - \{z_2\} \gg \\ \nu_2 := \ll Dz_2 - \{z_3\} \gg \\ \nu_3 := \ll Dz_3 - \{z_4\} \gg \\ \\ \vdots \\ \\ \nu_{r-1} := \ll Dz_{r-1} - \{z_r\} \gg \\ \nu_r := \ll Dz_r - CA^r W^{-1}\{z\} \gg (= CA^{r-1}B \ll \lambda \gg) \end{array}\right. \tag{27}$$

and

$$\lambda = (CA^{r-1}B)^{-1}[D^{(r-1)}\nu_1 + \ldots + D\nu_{r-1}] + \nu_r. \tag{28}$$

Each distribution ν_i is of the form

$$\langle \nu_i, \varphi \rangle = \int_I \varphi \, d\nu_i, \ \ \forall \varphi \in C_0^\infty(I) \tag{29}$$

where $d\nu_i$ is a Stieltjes measure generated by a $\mathcal{F}_\infty(\tilde{I}; \mathbb{R})$-function.

Let us now denote $d\{z_i\}$ $(1 \leq i \leq r)$ the Stieltjes measures defined by

$$\langle \ll Dz_i \gg, \varphi \rangle = \int_I \varphi \, d\{z_i\}, \ \ \forall \varphi \in C_0^\infty(I), \tag{30}$$

that is also the Stieltjes measure generated by $\{z_i\}$. We see that

$$d\nu_i = d \ll Dz_i - \{z_{i+1}\} \gg = d\{z_i\} - \{z_{i+1}\}(t)dt, \ \ (1 \leq i \leq r-1) \tag{31}$$

and

$$d\nu_r = d \ll Dz_r - CA^r W^{-1}\{z\} \gg = d\{z_r\} - CA^r W^{-1}\{z\}(t)dt. \tag{32}$$

If K denotes a closed convex cone then the expression $d\nu_i \in K$ has a sense that has been specified in Section 2 (see (2) and (3)). The constraints will be specified and discussed in the following section.

Remark 2 *i) The structure of the system in (26) ensures that necessarily z_i is a distribution of degree $\leq i$ while λ is a distribution of degree $\leq r+1$. It results that $z_1 \equiv \{z_1\}$ and the zero-dynamics is an ODE, i.e.*

$$\dot{\xi}(t) = A_\xi \xi(t) + B_\xi z_1(t), \forall t \in I.$$

ii) The mapping $\xi(.)$ is continuous on $\mathbb{R}^+$. Indeed, suppose that $t_n \to t^*$. We have

$$\|\xi(t_n) - \xi(t^*)\| \leq |\int_{t_n}^{t^*} (A_\xi \xi(s) + B_\xi z_1(s))ds|.$$

The mappings $\xi(.), z_1(.)$ are BV (and thus bounded) on compact subsets of $\mathbb{R}^+$ and thus there exists a constant $C > 0$ such that (for n large enough):

$$\|\xi(t_n) - \xi(t^*)\| \leq C|t_n - t^*|.$$

It results that $\xi(t_n) \to \xi(t^*)$.

Measure Differential Formalism. The Distributional Formalism contains all the information on the dynamics. However, it would be convenient to end up with a "weaker" formalism for (23) (24) by requiring that the solutions z_i of the system in (23) (24) are regular distributions z_i generated by right continuous functions of special locally bounded variation. More precisely, our problem consists to find $z_1, ..., z_r, \xi_1, ..., \xi_{n-r} \in \mathcal{F}_\infty(\mathbb{R}^+;\mathbb{R})$ such that

$$\left\{\begin{array}{l} dz_1 = z_2(t)dt + d\nu_1 \\ dz_2 = z_3(t)dt + d\nu_2 \\ dz_3 = z_4(t)dt + d\nu_3 \\ \\ \vdots \\ dz_i = z_{i+1}(t)dt + d\nu_i \\ \vdots \\ dz_{r-1} = z_r(t)dt + d\nu_{r-1} \\ dz_r = CA^r W^{-1} z(t)dt + CA^{r-1}Bd\nu_r \\ \\ \dot{\xi}(t) = A_\xi \xi(t) + B_\xi z_1(t) \end{array}\right. \tag{33}$$

where $d\nu_i$ denotes Radon measure and dz_i is the Stieltjes measure generated by z_i (see Section 2). If K denotes a closed convex cone then the expression $d\nu_i \in K$ can be defined as in Section 2.

Here, all distributions, let us say T, invoked in the system (23) (24) are of degree ≤ 1 and thus $\ll T \gg = T$. Here, we have also $\{z_i\} \equiv z_i$ $(1 \leq i \leq r)$.

Remark 3 *In Mechanics one has $r = 2$ and λ is of degree 2 because $\nu_1 = 0$ (the position $z_1(\cdot)$ is absolutely continuous, see e.g. [45, 70]). In dissipative electrical circuits with $r = 1$ then possible inconsistent initial data on z_1 may lead to $\nu_1(= \lambda)$ of degree 2 [18]. The measures ν_i and the distribution λ in (24) play a similar role to the Lagrange multiplier in Mechanics with unilateral contact. The idea of the extended Moreau's sweeping process is to represent these atomic distributions in a way similar to what is done in Mechanics and to subsequently take advantage of this formalism to derive a time-stepping numerical algorithm. Moreover, viewing the dynamics as an equality of distributions as in (24) paves the way towards time-discretization with time-stepping algorithms, i.e. numerical schemes working without event detection procedures and with constant time-step.*

Remark 4 *The idea of observing the derivatives of the variable $w(\cdot)$ in order to analyse unilaterally constrained systems is certainly not new, and has been often used previously [30, 72]. Also it has been long well known in DAE theory that higher degree distributions can occur due to inconsistent initial state [39, 21]. Therefore introducing such ingredients in dynamical systems is by far not new. It is however clear that the gap between DAE and complementarity systems, is not trivial.*

4 The extended Moreau's sweeping process

4.1 Preliminaries

Let us first recall that in order to simplify the presentation we shall continue to assume in many places that $m = 1$.

Let Φ be a nonempty closed convex subset of $\mathbb{R}$. We denote by $T_\Phi(x)$ the tangent cone of Φ at $x \in \mathbb{R}$ defined by

$$T_\Phi(x) = \overline{\text{cone}}(\Phi - \{x\}) \tag{34}$$

where $\text{cone}(\Phi - \{x\})$ denotes the cone generated by $\Phi - \{x\}$. The definition in (34) allows us to take into account constraints violations. Note that

$$T_{\mathbb{R}^+}(x) = \begin{cases} \mathbb{R} & \text{if} \quad x > 0 \\ \mathbb{R}^+ & \text{if} \quad x \leq 0 \end{cases}$$

and

$$T_{\mathbb{R}}(x) = \mathbb{R}.$$

Let us now set

$$\Phi := \mathbb{R}^+$$

For $z \in \mathbb{R}^r$, we set

$$Z_i = (z_1, z_2, ..., z_i), \quad (1 \le i \le r). \tag{35}$$

We define

$$T^0_\Phi(Z_1) = \Phi, \;\; T^1_\Phi(Z_1) = T_\Phi(z_1), \;\; T^2_\Phi(Z_2) = T_{T^1_\Phi(Z_1)}(z_2),$$
$$T^i_\Phi(Z_i) = T_{T^{i-1}_\Phi(Z_{i-1})}(z_i).$$

Remark 5 *If $m \ge 2$ then $\Phi = (\mathbb{R}^+)^m$ and*

$$T^i_\Phi(Z_i) = \times^m_{l=1} T^i_\Phi(Z^l_i).$$

We shall keep the notation Φ noting that in general when one starts from the ZD dynamics, one gets $\Phi = \mathbb{R}^+$. Moreover we also keep in mind the extension of the material that follows towards formalisms involving time-dependent convex sets $\Phi(t)$ and not necessarily being the ZD dynamics of a given system.

Starting from (23), (24) the extended sweeping process is written as follows

$$d\nu_i \in -\partial\psi_{T^{i-1}_\Phi(\{Z_{i-1}\}(t^-))}(\{z_i\}(t^+)) \quad \text{on } \tilde{I}, \quad (1 \le i \le r) \tag{36}$$

with $d\nu_i$ in (29)–(32). The sets $\partial\psi_{T^{i-1}_\Phi(\{Z_{i-1}\}(t^-))}(\{z_i\}(t^+))$ are nonempty closed convex cones and therefore the inclusions in (36) make sense (see Section 2). Recall that $I =]0, +\infty[$ and $\tilde{I} = [0, +\infty[$.

Remark 6 *i) Recall that if $z_i \in T^{i-1}_\Phi(Z_{i-1})$ then*

$$\partial\psi_{T^{i-1}_\Phi(Z_{i-1})}(z_i) = \{w \in \mathbb{R} : \langle w, v - z_i\rangle \le 0, \forall v \in T^{i-1}_\Phi(Z_{i-1})\}$$

is the outward normal cone to $T^{i-1}_\Phi(z_1, ..., z_{i-1})$ at z_i.
ii) Note that

$$T^{i-1}_\Phi(Z_{i-1}) = \mathbb{R} \;\; \Rightarrow \;\; \partial\psi_{T^{i-1}_\Phi(Z_{i-1})}(z_i) = \{0\},$$
$$T^{i-1}_\Phi(Z_{i-1}) = \mathbb{R}^+ \;\text{ and }\; z_i > 0 \;\; \Rightarrow \partial\psi_{T^{i-1}_\Phi(Z_{i-1})}(z_i) = \{0\},$$
$$T^{i-1}_\Phi(Z_{i-1}) = \mathbb{R}^+ \;\text{ and }\; z_i \le 0 \;\; \Rightarrow \partial\psi_{T^{i-1}_\Phi(Z_{i-1})}(z_i) = \mathbb{R}^-.$$

Starting from (16) and (18) one is tempted to write the inclusion

$$Dz_r - CA^r W^{-1} z(t) \in -CA^{r-1}B\, \partial\psi_\Phi(z_1(t^+)) \quad \text{on } \tilde{I},$$

which makes sense only if λ is a measure since $\partial\psi_\Phi(z_1(\cdot^+))$ is a cone. This inclusion is replaced by

$$d\nu_r \in -\, \partial\psi_{T_\Phi^{r-1}(\{Z_{r-1}\}(t^-))}(\{z_r\}(t^+)) \quad \text{on } \tilde{I},$$

in (36). The positivity of λ is now understood as the positivity of ν_r (see also Theorem 1 below).

Remark 7 *It is then important to see that the distributional dynamics*

$$D^r z_1 = Dz_r = CA^r W^{-1}\{z\} + CA^{r-1}B\lambda \tag{37}$$

with λ in (36) (25), is equivalent to (24) (36). Notice that (25) (36) (37) is the same as (23) (25) (36).

Definition 1 (Higher Order Moreau's Sweeping Process) *The higher order (or extended) sweeping process is the dynamical system represented in (23)(24) (25) (36).*

Remark 8 *It is noteworthy that though the formalism is presented in the coordinates in (18), its implementation requires the knowledge of A, B, C and $x(\cdot)$ only. We however believe that the state space representation in (18) allows one to better understand the underlying dynamics.*

4.2 Mathematical Formalisms

Let us now complete the formalisms introduced in Section 3.2.

Distributional Formalism. Using (26), (27), (28), (29), (30), (31), (32) and (36), our mathematical problem reads: Find $z_1, ..., z_r \in \mathcal{T}_{r-1}(I)$ and $\xi_i \in \mathcal{T}_{r-1}(I)$ $(1 \le i \le n-r)$ satisfying the distributional equations

$$\left\{ \begin{array}{l} Dz_1 - z_2 = 0 \\ Dz_2 - z_3 = 0 \\ Dz_3 - z_4 = 0 \\ \\ \vdots \\ \\ Dz_{r-1} - z_r = 0 \\ Dz_r - CA^r W^{-1} z = CA^{r-1}B\lambda \\ \\ D\xi = A_\xi \xi + B_\xi z_1 \end{array} \right. \tag{38}$$

$$\lambda = (CA^{r-1}B)^{-1}[D^{(r-1)} \ll Dz_1 - \{z_2\} \gg + \ldots +$$

$$+D \ll Dz_{r-1} - \{z_r\} \gg] + \ll Dz_r - CA^r W^{-1}\{z\} \gg \tag{39}$$

and

$$\begin{cases} d\{z_1\} - \{z_2\}(t)dt \in -\partial\psi_{T^0_\Phi}(\{Z_1\}(t^+)), \\ \vdots \\ d\{z_i\} - \{z_{i+1}\}(t)dt \in -\partial\psi_{T^{i-1}_\Phi(\{Z_{i-1}\}(t^-))}(\{z_i\}(t^+)), \\ \vdots \\ d\{z_{r-1}\} - \{z_r\}(t)dt \in -\partial\psi_{T^{r-2}_\Phi(\{Z_{r-2}\}(t^-))}(\{z_{r-1}\}(t^+)), \\ (CA^{r-1}B)^{-1}[d\{z_r\} - CA^r W^{-1}\{z\}(t)dt] \in -\partial\psi_{T^{r-1}_\Phi(\{Z_{r-1}\}(t^-))}(\{z_r\}(t^+)). \end{cases} \tag{40}$$

The relations given in (40) are formulated on $\tilde{I}$ and have to be interpreted in the following sense: Find nonnegative real-valued Radon measures $d\mu_i \quad (1 \leq i \leq r)$ relative to which the Lebesgue measure dt and the Stieltjes measure $d\{z_i\}$ possess densities $\frac{dt}{d\mu_i}$ and $\frac{d\{z_i\}}{d\mu_i}$ respectively such that:

$$\frac{d\{z_i\}}{d\mu_i}(t) - \{z_{i+1}\}(t)\frac{dt}{d\mu_i}(t) \in -\partial\psi_{T^{i-1}_\Phi(\{Z_{i-1}\}(t^-))}(\{z_i\}(t^+)),$$

$$d\mu_i - \text{a.e. } \; t \in \tilde{I} \qquad (1 \leq i \leq r-1) \tag{41}$$

and

$$(CA^{r-1}B)^{-1}[\frac{d\{z_r\}}{d\mu_r}(t) - CA^r W^{-1}\{z\}(t)\frac{dt}{d\mu_r}(t)] \in -\partial\psi_{T^{r-1}_\Phi(\{Z_{r-1}\}(t^-))}(\{z_r\}(t^+)),$$

$$d\mu_r - \text{a.e. } \; t \in \tilde{I}. \tag{42}$$

By convention, we set

$$z(0^-) = z_0 \tag{43}$$

where z_0 is given in $\mathbb{R}^n$ so as to define some appropriate initial condition.

Measure Differential Formalism. Using (33) and (36), our mathematical problem reads: Find $z_i \in \mathcal{F}_\infty(\tilde{I};\mathbb{R})$ $(1 \leq i \leq r)$ and $\xi_i \in \mathcal{F}_\infty(\tilde{I};\mathbb{R})(1 \leq i \leq n-r)$ such that

$$dz_i - z_{i+1}(t)dt \in -\partial\psi_{T^{i-1}_\Phi(z_{i-1}(t^-))}(z_i(t^+)) \quad \text{on} \quad \tilde{I} \;\; (1 \leq i \leq r-1) \tag{44}$$

$$(CA^{r-1}B)^{-1}[dz_r - CA^r W^{-1}z(t)dt] \in -\partial\psi_{T^{r-1}_\Phi(z_{r-1}(t^-))}(z_r(t^+)) \quad \text{on} \quad \tilde{I} \tag{45}$$

and

$$\dot{\xi}(t) - A_\xi \xi(t) - B_\xi z_1(t) = 0, \text{ a.e. } t \in \tilde{I} \tag{46}$$

The system in (44) and (45) has to be interpreted in the following sense: Find nonnegative real-valued Radon measure $d\mu_i$ relative to which the Lebesgue measure dt and the Stieltjes measure dz_i possess densities $\frac{dt}{d\mu_i}$ and $\frac{dz_i}{d\mu_i}$ respectively such that

$$\frac{dz_i}{d\mu_i}(t) - z_{i+1}(t)\frac{dt}{d\mu_i}(t) \in -\partial\psi_{T_\Phi^{i-1}(z_{i-1}(t^-))}(z_i(t^+)), \;\; d\mu_i - \text{a.e. } t \in \tilde{I} \;\; (1 \le i \le r-1) \tag{47}$$

and

$$(CA^{r-1}B)^{-1}[\frac{dz_r}{d\mu_r}(t) - CA^r W^{-1} z(t)\frac{dt}{d\mu_r}(t)] \in -\partial\psi_{T_\Phi^{r-1}(Z_{r-1}(t^-))}(z_r(t^+)),$$

$$d\mu_r - \text{a.e. } t \in \mathbb{R}. \tag{48}$$

Remark 9 *i) The model in (47)-(48) is the same that the one given in (41)-(42) since here* $\{z_i\} \equiv z_i$.
ii) If z_1 *is piecewise continuous then the general solution* ξ *of (46) is given by*

$$\xi(t) = e^{A_\xi t}D + \int_0^t e^{A_\xi(t-\tau)}B_\xi z_1(\tau)d\tau,$$

where $D \in \mathbb{R}^{n-r}$.

Remark 10 *It is noteworthy that the differential inclusions which are considered in this paper, are specific DI (which could be named* unbounded *DI), which cannot be analysed using the tools for "standard" DIs of the form* $\dot{x}(t) \in F(x(t))$ *as in [5]. Especially the basic assumptions for standard DIs are that* $F(x)$ *is compact for each* x*, and that a linear growth condition holds. Such assumptions do not hold for the inclusions considered here, since the sets in the right hand sides are normal cones, hence unbounded sets. Such a kind of differential inclusions gave rise to all the mathematical and numerical studies on the sweeping process and mechanical systems subject to unilateral constraints [48, 49, 50, 51, 20, 7, 34, 45, 19, 38, 35, 36, 37, 6, 63, 64, 30, 43, 70, 15].*

Variational Inequalities. The system in (41)-(42) (and consequently the one in (47)-(48) too) together with the zero-dynamics can be written as the evolution variational inequalities:

$$\begin{cases} \langle \frac{d\{z_i\}}{d\mu_i}(t) - \{z_{i+1}\}(t)\frac{dt}{d\mu_i}(t), v - \{z_i\}(t^+)\rangle \geq 0, \;\; \forall \;\; v \in T_\Phi^{i-1}(\{Z_{i-1}\}(t^-)), \\ \\ d\mu_i - \text{a.e.} \;\; t \in \tilde{I} \;\; (1 \leq i \leq r-1) \\ \\ \langle (CA^{r-1}B)^{-1}[\frac{d\{z_r\}}{d\mu^{w\lambda}}(t) - CA^r W^{-1}\{z\}(t)\frac{dt}{d\mu_r}(t)], v - \{z_r\}(t^+)\rangle \geq 0, \\ \\ \forall \;\; v \in T_\Phi^{r-1}(\{Z_{r-1}\}(t^-)), \;\; d\mu_r - \text{a.e.} \;\; t \in \tilde{I} \\ \\ \\ \langle \dot{\xi}(t) - A_\xi \xi(t) - B_\xi z_1(t), v - \xi(t)\rangle \geq 0, \;\; \forall \, v \in \mathbb{R}, \;\; \text{a.e.} \;\; t \in \tilde{I}. \end{cases} \tag{49}$$

The following Proposition shows the link between the distributional formalism of the measure differential formalism.

Proposition 2 *i) Let* $(z_1, ..., z_r, \xi) \in (\mathcal{T}_{r-1}(I)^n$ *be a solution of Problem (38) (39) (40) (43). Then* $z_1 = \{z_1\} \in \mathcal{F}_\infty(\tilde{I}; \mathbb{R})$, $z_i \in \mathcal{T}_{i-1}(I)\,(2 \leq i \leq r), \xi = \{\xi\} \in (\mathcal{F}_\infty(\tilde{I}; \mathbb{R}))^{n-r}$ *and* $(\{z_1\}, ..., \{z_r\}, \xi)$ *is a solution of Problem (44) (45) (46) (43).*

ii) Let $(w_1, ..., w_r, \xi) \in (\mathcal{F}_\infty(\tilde{I}; \mathbb{R}))^n$ *be a solution of Problem (44) (45) (46) (43). Let* $(z_1, ..., z_r, \xi)$ *be defined by*

$$z_1 := w_1$$

and

$$z_i := w_i + \sum_{j=1}^{i-1} \Big(\sum_{t_k \in E_0(w_j)} (w_j(t_k^+) - w_i(t_k^-))\delta_{t_k}^{(i-j-1)} \Big) \quad (2 \leq i \leq r),$$

Then $(z_1, ..., z_r, \xi) \in (\mathcal{T}_{r-1}(I)^n$ *and is a solution of Problem (38) (39) (40) (43).* ■

4.3 The State Jump Mapping

In this Section, we examine some properties of the solution of problem (38) (39) (40) (43) (and consequently of problem (44) (45) (46) (43)). The existence of such a solution is thus assumed for the time being. Let us first remark that the following is true

Lemma 1 *Let* $z_1, \ldots, z_r \in \mathbb{R}$ *be given and let* $\sigma_1, \ldots, \sigma_r \in \mathbb{R}$ *such that*

$$\sigma_i \in -\partial\psi_{T_\Phi^{i-1}(Z_{i-1})}(z_i) \;\; \textit{for all} \;\; 1 \leq i \leq r.$$

Then the inclusion

$$\partial\psi_{T_\Phi^{r-1}(Z_{r-1})}(z_r) \subseteq \partial\psi_\Phi(z_1)$$

holds. Moreover,

$$\begin{cases} z_1 > 0 \Longrightarrow \sigma_r = 0 \\ z_1 = 0 \Longrightarrow \sigma_r \geq 0. \end{cases}$$

Proof: If $z_1 > 0$ then $T_\Phi(z_1) = \mathbb{R}$ and $\partial\psi_\Phi(z_1) = \{0\}$, and since $T_\Phi^{i-1}(Z_{i-1}) = \mathbb{R}$ then $\partial\psi_{T_\Phi^{i-1}(Z_{i-1})}(z_i) = \{0\}$ for all $1 \le i \le r$. In particular we deduce that $\sigma_r = 0$. If $z_1 = 0$ then $\partial\psi_{\mathbb{R}^+}(z_1) = \mathbb{R}^-$. Depending on the values of $z_2, \ldots, z_i$ being positive or non-positive, one may have $\partial\psi_{T_\Phi^i(Z_i)}(z_{i+1}) = \mathbb{R}^-$ or $\partial\psi_{T_\Phi^i(Z_i)}(z_{i+1}) = \{0\}$ for all $1 \le i \le r-1$. Indeed assume that $z_1 = z_2 = \ldots = z_j = 0$ and $z_{j+1} > 0$ (this implies that $z \succeq 0$). Then $\Phi = T_\Phi^0(Z_1) = T_\Phi^1(Z_1) = T_\Phi^2(Z_2) = T_\Phi^3(Z_3) = \ldots = T_\Phi^j(Z_j) = \mathbb{R}^+$ and $T_\Phi^{j+1}(Z_{j+1}) = T_\Phi^{j+2}(Z_{j+2}) = \ldots = T_\Phi^{r^{r\lambda}-1}(Z_{r-1}) = \mathbb{R}$. This can be seen since for instance $T^{j+2}(Z_{j+2}) = T_{T_\Phi^{j+1}(Z_{j+1})}(z_{j+2}) = T_{\mathbb{R}}(z_{j+2}) = \mathbb{R}$ because also $T_\Phi^{j+1}(Z_{j+1}) = T_{\mathbb{R}^+}(z_{j+1}) = \mathbb{R}$. Consequently $\partial\psi_{T_\Phi^i(Z_i)}(z_{i+1}) = \mathbb{R}^-$ for all $0 \le i \le j$, whereas $\partial\psi_{T_\Phi^i(Z_i)}(z_{i+1}) = \{0\}$ for $j+1 \le i \le r-1$. We conclude that under such conditions $\sigma_i \ge 0$ for all $1 \le i \le j+1$, and $\sigma_i = 0$ for all $j+2 \le i \le r$. Consequently $\sigma_r \ge 0$ when $z_1 = 0$. The inclusion is also proved. ■

Let $t \in \tilde{I}$ be given. A direct consequence of the higher order sweeping process appears as soon as one measures the set $\{t\}$ with the Radon measure $d\nu_i$. Indeed, as seen in Section 2, we get from (36):

$$d\nu_i(\{t\}) \in -\partial\psi_{T_\Phi^{i-1}(\{Z_{i-1}\}(t^-))}(\{z_i\}(t^+)). \tag{50}$$

Then

$$d\nu_i(\{t\}) = d\{z_i\}(\{t\}) - \{z_{i+1}\}(t)dt(\{t\}) = d\{z_i\}(\{t\}) = \{z_i\}(t^+) - \{z_i\}(t^-)$$

for all $1 \le i \le r-1$, and

$$\begin{aligned}(CA^{r-1}B)d\nu_r(\{t\}) &= d\{z_r\}(\{t\}) - CA^rW^{-1}\{z\}(t)dt(\{t\}) \\ &= d\{z_r\}(\{t\}) = \{z_r\}(t^+) - \{z_r\}(t^-).\end{aligned}$$

Obviously, the same result holds in the framework of the measure differential formalism (in this case $z_i \equiv \{z_i\}$).

These results ensure that the extended sweeping process inclusion defines a well-posed state jump mapping and the following holds.

Proposition 3 *Let $t \in \tilde{I}$ be given. We have*

$$\{z_i\}(t^+) - \{z_i\}(t^-) \in -\partial\psi_{T_\Phi^{i-1}(\{Z_{i-1}\}(t^-))}(\{z_i\}(t^+)), \ \ (1 \le i \le r-1) \tag{51}$$

and

$$\{z_r\}(t^+) - \{z_r\}(t^-) \in -CA^{r-1}B\ \partial\psi_{T_\Phi^{r-1}(\{Z_{r-1}\}(t^-))}(\{z_r\}(t^+)). \tag{52}$$

■

Another direct consequence of (36) is that

$$\{z_i\}(t^+) \in T_\Phi^{i-1}(\{Z_{i-1}\}(t^-)), \ \ (1 \le i \le r_{w\lambda}). \tag{53}$$

Our convention in (43) together with (50) yield

$$\{z_i\}(0^+) \in z_{0,i} - \partial\psi_{T_\Phi^{i-1}(Z_{0,i-1})}(\{z_i\}(0^+)),\ \ (1 \le i \le r-1)$$

and

$$\{z_r\}(0^+) \in z_{0,r} - CA^{r-1}B\,\partial\psi_{T_\Phi^{r-1}(Z_{0,r-1})}(\{z_r\}(0^+)).$$

where $z_{0,i}$ is the ith-component of z_0 and $Z_{0,i} = (z_{0,1}, \dots, z_{0,i})$.

Theorem 1 *Let $t \in \tilde{I}$ be given. Then*

$$0 \le z_1(t^+) \perp d\nu_r(\{t\}) \ge 0 \tag{54}$$

Proof: Let us first recall (see Remark 2) that $z_1 = \{z_1\}$. Moreover $z_1(t^+) \in \Phi$ and thus $z_1(t^+) \ge 0$. From Lemma 1, we obtain that if $z_1(t^+) > 0$ then $d\nu_r(\{t\}) = 0$ while if $z_1(t^+) = 0$ then $d\nu_r(\{t\}) \ge 0$. We deduce that if $d\nu_r(\{t\}) > 0$ then necessarily $z_1(t^+) = 0$. The result follows. ■

We may write $d\nu_i$ as

$$d\nu_i = \chi_i(t)dt + d\mathcal{J}_i, \tag{55}$$

where $\chi_i \in \mathcal{F}_\infty(\in \tilde{I}; \mathbb{R})$ and $d\mathcal{J}_i$ is an atomic measure with countable set of atoms generated by a right continuous jump function $\mathcal{J}_i$. The following holds.

Proposition 4 *We have*

$$\chi_i(t) = 0, \ \text{ a.e. } \ t \in \tilde{I}, \ \ (1 \le i \le r-1), \tag{56}$$

$$\chi_r(t) \in -\partial\psi_{T_\Phi^{r-1}(\{z_1\}(t^-),\dots,\{z_{r-1}\}(t^-))}(\{z_r\}(t^+)), \ \text{ a.e. } \ t \in \tilde{I}, \tag{57}$$

and

$$0 \le z_1(t^+) \perp \chi_r(t) \ge 0, \ \text{ a.e. } \ t \in \tilde{I}. \tag{58}$$

Proof: Let $1 \le i \le r-1$ be given. We know that $Dz_i = z_{i+1}$ and thus $\{Dz_i\} = \{z_{i+1}\}$. Thus $\nu_i = \ll Dz_i - \{Dz_i\} \gg$. It results that $d\nu_i$ is an atomic measure and thus $\chi_i(t) = 0,$ a.e. $t \in \tilde{I}$ and (56) holds. The result in (57) is a direct consequence of [71, Theorem 5]. The complementarity result in (58) can be proved as in Theorem 1 thanks to (57) and Lemma 1. ■

Theorem 1 implies that on intervals $[\tau - \epsilon, \tau)$, $\epsilon > 0$, on which $z_1(t) = 0$, there exists a multiplier $\lambda(t) := d\nu_r(\{t\})$ that belongs to $-\partial\psi_{\mathbb{R}^+}(0) = \mathbb{R}^+$, equivalently which satisfies $0 \le z_1(t^+) \perp \lambda(t) \ge 0$ and is the solution of the LCP in (59). Equivalently, $d\nu_r$ is the sum of an atomic measure $d\mathcal{J}_r$, $\text{supp}(d\mathcal{J}_r) \subset \{t \in \mathbb{R}^+ : z_1(t) = 0\}$, that corresponds to jumps in $z_r(\cdot)$, and of a Lebesgue measure $\chi_r(t)dt$ where $\chi_r(t)$ is the solution of the LCP in (59).

Remark 11 *From (57) it follows that the following LCP holds on intervals* $[\tau, \tau+\epsilon)$, $\epsilon > 0$, *on which* $z_1(t) = 0$:

$$0 \leq \lambda(t) \perp CA^r W^{-1}\{z\}(t^+) + CA^{r-1}B\lambda(t) \geq 0 \quad (t \geq 0) \tag{59}$$

where $\lambda(t) = \chi_r(t)$ *on* $[\tau, \tau+\epsilon)$. *This LCP monitors the evolution of the multiplier for* $t \in [\tau, \tau+\epsilon)$. *The matrix of this LCP is the leading Markov parameter of the transfer function* $\lambda \mapsto w$, *i.e.* $CA^{r-1}B$. *If* $CA^{r-1}B$ *is a P-matrix [58], then the LCP in (59) has a unique solution for any* $CA^r W^{-1}\{z\}(t^+)$. *One sees that when discretizing the continuous dynamics, one will take advantage of the ZD representation to construct an LCP to compute* $\lambda(t)$ *at each step.*

The following two results give some equivalent characterizations of the relations given in Proposition 3. Their proofs are straightforward. These results are here presented for the general case $m \geq 1$.

Proposition 5 *Let* $m \geq 1$. *The following equivalences hold for all* $1 \leq i \leq r-1$ *and all* $t \in \tilde{I}$:

$$\begin{array}{lr} \{z_i\}(t^+) - \{z_i\}(t^-) \in -\partial\psi_{T_\Phi^{i-1}(\{Z_{i-1}\}(t^-))}(\{z_i\}(t^+)) & \mathbf{(a)} \\ \Updownarrow & \\ \{z_i\}(t^+) = prox\left[T_\Phi^{i-1}(\{Z_{i-1}\}(t^-)); \{z_i\}(t^-)\right] & \mathbf{(b)} \end{array} \tag{60}$$

■

This shows that jumps are automatically taken into account by the dynamics as it is written in (40) (or in (44)-(45)). We notice also that the lower triangular structure of the tangent cones which appear in (36) merely reflects the way the measures $d\nu_i$ appear in (40) (or in (44)-(45)).

Proposition 6 *Let* $m \geq 1$. *Suppose that the matrix* $CA^{r-1}B$ *is symmetric and positive definite. Then for all* $t \in \tilde{I}$:

$$\begin{array}{lr} \{z_r\}(t^+) - \{z_r\}(t^-) \in -CA^{r-1}B\ \partial\psi_{T_\Phi^{r-1}(\{Z_{r-1}\}(t^-))}(\{z_r\}(t^+)) & \mathbf{(a)} \\ \Updownarrow & \\ \{z_r\}(t^+) = prox_{(CA^{r-1}B)^{-1}}\left[T_\Phi^{r-1}(\{Z_{r-1}\}(t^-)); \{z_r\}(t^-)\right] & \mathbf{(b)} \end{array} \tag{61}$$

■

Remark 12 *We note that from Propositions 5 and 6, the measures $d\nu_i$ are uniquely determined at jump instants. The formalism in (36) corresponds to some kind of "plastic" impacts since the high-order inconsistent derivatives jump to zero. It is quite possible to incorporate different jumps by modifying the right-hand sides and changing the argument in the subdifferentials from $\{z_i\}(t^+)$ to $\frac{\{z_i\}(t^+)+e_i\{z_i\}(t^-)}{1+e_i}$ [43, 54]. Then (60) becomes*

$$\begin{array}{lr} \{z_{i+1}\}(t^+) - \{z_{i+1}\}(t^-) \in -\partial\psi_{T_\Phi^i(\{Z_i\}(t^-))}\left(\frac{\{z_{i+1}\}(t^+)+e_{i+1}\{z_{i+1}\}(t^-)}{1+e_{i+1}}\right) & (\mathbf{a}) \\ \Updownarrow & \\ \{z_{i+1}\}(t^+) = -e_{i+1}\{z_{i+1}\}(t^-) + (1+e_{i+1})\mathit{prox}\left[T_\Phi^i(\{Z_i\}(t^-)); \{z_{i+1}\}(t^-)\right] & (\mathbf{b}) \end{array} \tag{62}$$

The choice of the coefficients e_i depends on the application. From (62) it follows that if $T_\Phi^i(\{Z_i\}(t^-)) = (\mathbb{R}^+)^m$ and if $\{z_{i+1}\}(t^-) < 0$, then $\{z_{i+1}\}(t^+) = -e_{i+1}\{z_{i+1}\}(t^-)$. If $m \geq 2$ and a uniform relative degree, then a similar operation can be performed. However in such a case (multi-constraint system), one may define mr coefficients $e_{i,j}$, where the i refers to the order of the derivative of $z_1^j(\cdot)$, while the j refers to the component of the m-vector $z_1 = (z_1^1, \ldots, z_1^m)^T$ (see (21)).

4.4 Dissipativity properties

Monotonicity and dissipativity are at the core of the second order sweeping process. They are crucial properties for well-posedness results [9, 6] and control/stability [12, 41, 33, 18]. The following results are consequently of interest.

Lemma 2 *Let $z_1, z_2, ..., z_r \in \mathbb{R}$ be given. The following inclusion holds*

$$\partial\psi_{T_\Phi^{i-1}(Z_{i-1})}(z_i) \subseteq \partial\psi_{T_\Phi^{i-2}(Z_{i-2})}(z_{i-1}), \tag{63}$$

for all $1 \leq i \leq r$.

Proof : Let $z_1 \leq z_2 \leq \ldots = z_j = 0$ and $z_{j+1} > 0$. Then as already shown in the proof of lemma 1, one has $\partial\psi_{T_\Phi^i(Z_i)}(z_{i+1}) = \mathbb{R}^-$ for all $0 \leq i \leq j$, and $\partial\psi_{T_\Phi^i(Z_i)}(z_{i+1}) = \{0\}$ for all $j+1 \leq i \leq r-1$. So one sees that in particular it always holds that $\partial\psi_{T_\Phi^k(Z_k)}(z_{k+1}) \subseteq \partial\psi_{T_\Phi^{k-1}(Z_{k-1})}(z_k)$ for any $1 \leq k \leq r-1$. ■

Let $z \in \mathbb{R}^r$ be given. The domain and the graph of the multivalued operator $x \rightarrow \partial\psi_{T_\Phi^{i-1}(Z_i)}(x)$ $(1 \leq i \leq r)$ are given respectively by

$$D(\partial\psi_{T_\Phi^{i-1}(Z_i)}) := \{x \in \mathbb{R} : \partial\psi_{T_\Phi^{i-1}(Z_i)}(x) \neq \emptyset\} = T_\Phi^{i-1}(Z_i)$$

and

$$\mathcal{G}(\partial\psi_{T_{\Phi}^{i-1}(Z_i)}) := \{(x,y) \in \mathbb{R}\times\mathbb{R} : x \in D(\partial\psi_{T_{\Phi}^{i-1}(Z_i)}) \text{ and } y \in \partial\psi_{T_{\Phi}^{i-1}(Z_i)}\}.$$

It is well-known that the graph of an operator which is defined as the subdifferential of a proper convex and lower semicontinuous function is maximal monotone (see e.g. [9]). Here the set $T_{\Phi}^{i-1}(Z_i)$ is nonempty closed and convex and the mapping $\psi_{T_{\Phi}^{i-1}(Z_i)}$ is thus proper convex and lower semicontinuous. It results that $\mathcal{G}(\partial\psi_{T_{\Phi}^{i-1}(Z_i)})$ is monotone, i.e.

$$\langle x_1 - x_2, y_1 - y_2\rangle \geq 0, \forall (x_1,y_1),(x_2,y_2) \in \mathcal{G}(\partial\psi_{T_{\Phi}^{i-1}(Z_i)})$$

and is not properly included in any monotone subset of $\mathbb{R}\times\mathbb{R}$.
Here we have

$$T_{\Phi}^{i-1}(\{Z_{i-1}\}(t^-)) \ni \{z_i\}(t^+) \perp -d\nu_i(\{t\}) \in \partial\psi_{T_{\Phi}^{i-1}(\{Z_{i-1}\}(t^-))}(\{z_i\}(t^+)). \tag{64}$$

For example, suppose that $T_{\Phi}^{i-1}(\{Z_{i-1}\}(t^-)) = \mathbb{R}^+$, $\forall t \in (T_1, T_2]$ for some $0 \leq T_1 < T_2$. Then as a consequence of (64), we get the dissipativity property

$$\langle \{z_i\}(t_1^+) - \{z_i\}(t_2^+), d\nu_i(\{t_1^+\}) - d\nu_i(\{t_2^+\})\rangle \leq 0, \quad \forall t_1, t_2 \in (T_1, T_2), \quad (1 \leq i \leq r).$$

We notice that from (64) we get

$$T_{\Phi}(\{z_1\}(t^-)) \ni \{z_2\}(t^+) \perp -d\nu_2(\{t\}) \in \partial\psi_{T_{\Phi}(\{Z_1\}(t^-))}(\{z_2\}(t^+)) \subseteq N_{\Phi}(\{z_1\}(t^+)) \tag{65}$$

which is quite similar to the cone CP of the second order sweeping process (see [54, 53] [12, Equ.(16)]). This together with the formalisms developed in sections 3.2, 4.1 and 4.2, imply that the framework proposed in this paper contains the second order sweeping process.

We now investigate the dissipativity properties of the differential operator of the sweeping process.
Let us denote

$$\bar{G} = \begin{pmatrix} G \\ 0_{(n-r)\times r} \end{pmatrix} \in \mathbb{R}^{n\times r}, \tag{66}$$

with

$$G = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & \dots & 0 & CA^{r-1}B \end{pmatrix} \in \mathbb{R}^{r\times r} \tag{67}$$

and $H = (I_r \;\; 0_{n-r}) \in \mathbb{R}^{r\times n}$. Notice that the operator $d\nu \mapsto \{\bar{z}\}$ has the transfer matrix defined from the triple $(WAW^{-1}, \bar{G}, H)$.

Let the triple $(WAW^{-1}, \bar{G}, H)$ be positive real and the corresponding realisation be minimal. The positive realness of $(WAW^{-1}, \bar{G}, H)$ is equivalent by the Kalman-Yakubovic-Popov lemma (see appendix C) to having $J\bar{G} = H^T$ and $JWAW^{-1} + W^{-T}A^TW^TJ$ semi-negative definite for some positive definite and symmetric matrix J. Since $J = J^T$ and since G is full rank, the first equality implies that

$$J = \begin{pmatrix} G^{-1} & 0_{r\times(n-r)} \\ 0_{(n-r)\times r} & J_\xi \end{pmatrix}, \tag{68}$$

and $J_\xi = J_\xi^T \in I\!R^{(n-r)\times(n-r)}$ is positive definite.

Proposition 7 (Dissipation inequality) *Let us consider the extended sweeping process in (38) (39) (40) (43) with $m = 1$. Let the triple $(WAW^{-1}, \bar{G}, H)$ be positive real and the corresponding realisation be minimal. Then $CA^{r-1}B > 0$ and*

$$\{z\}(t_2^+)^TJ\{z\}(t_2^+) \leq \{z\}(t_1^-)^TJ\{z\}(t_1^-), \;\; \forall t_2 \geq t_1 \geq 0, \tag{69}$$

where J is given in (68).

Proof: The first statement of the proposition is a direct consequence of the positive definiteness of J.
Let us now set

$$d\nu := (d\nu_1, ..., d\nu_r, 0_{1\times(n-r)})^T.$$

We have

$$d\{z\} = WAW^{-1}\{z\}dt + \bar{G}d\nu = WAW^{-1}\{z\}^+dt + \bar{G}d\nu \tag{70}$$

and therefore

$$Jd\{z\} = JWAW^{-1}\{z\}^+dt + J\bar{G}d\nu = JWAW^{-1}\{z\}^+dt + H^Td\nu.$$

It results that

$$(\{z\}^+)^TJd\{z\} = (\{z\}^+)^T(JWAW^{-1}\{z\}^+)dt + (\{z\}^+)^TH^Td\nu.$$

Here $(\{z\}^+)^TH^Td\nu = \sum_{i=1}^r \{z_i\}^+\frac{d\nu_i}{d\mu_i}d\mu_i$ and thus $(\{z\}^+)^TH^Td\nu \equiv 0$ on $\mathbb{R}^+$ since the conditions in (40) (see also (41) and (42)) entail that $\{z_i\}(t^+)\frac{d\nu_i}{d\mu_i}(t) = 0$ for $d\mu_i$-a.e. $t \geq 0$. Consequently

$$(\{z\}^+)^TJd\{z\} = (\{z\}^+)^T(JWAW^{-1}\{z\}^+)dt.$$

We have

$$d(\{z\}^TJ\{z\}) = \sum_{i=1}^{r-1} d\{z_i\}^2 + (CA^{r-1}B)^{-1}d\{z_r\}^2 + d(\xi^TJ_\xi\xi)$$

and then using (8) and Remark 2 (ii) and recalling that here $CA^{r-1}B > 0$, $J_\xi = J_\xi^T$, J_ξ is positive definite, we see that

$$d(\{z\}^T J\{z\}) \leq \sum_{i=1}^{r-1} 2\{z_i\}^+ d\{z_i\} + 2(CA^{r-1}B)^{-1}\{z_r\}^+ d\{z_r\} + 2\xi^T J_\xi d\xi = 2(\{z\}^+)^T J d\{z\}$$

and thus

$$d(\{z\}^T J\{z\}) \leq 2(\{z\}^+)^T J d\{z\} = 2(\{z\}^+)^T (JWAW^{-1}\{z\}^+)dt =$$

$$= (\{z\}^+)^T (JWAW^{-1} + W^{-T}A^T W^T J)\{z\}^+ dt.$$

Then for $0 \leq t_1 \leq t_2$, we get

$$\{z\}^T(t_2^+)J\{z\}(t_2^+) - \{z\}(t_1^-)^T J\{z\}(t_1^-) = d(\{z\}^T J\{z\})([t_1, t_2]) =$$

$$= \int_{t_1}^{t_2} \{z\}(s^+)^T (JWAW^{-1} + W^{-T}A^T W^T J)\{z\}(s^+)ds \leq 0.$$

■

Proposition 7 and its proof are similar to [6, Proposition 7] and [12, Lemma 3] (see also [18, Theorem 11.2]), which apply to Lagrangian systems or to the case $r = 1$.

4.5 Well-Posedness Results

Let us first state some additional properties of the solution of problem (38) (39) (40) (43). Recall that for $z \in \mathbb{R}^n$, we use the notation $z^T = (\bar{z}^T, \xi^T)$ as in (17) with $\bar{z} \in \mathbb{R}^r$ and $\xi \in \mathbb{R}^{n-r}$.

Proposition 8 *Suppose that $CA^{r-1}B > 0$. Let $t \geq 0$ be given. i) If $\{z_1\}(t^-) > 0$ then the function $\{z_i\}(.)$ $(1 \leq i \leq r)$ is continuous at t . ii) If $\{z_1\}(t^-) \leq 0$ and $\{z_2\}(t^-) > 0$ then $\{z_1\}(t^+) = 0$ and the function $\{z_i\}(.)$ $(2 \leq i \leq r)$ is continuous at t. iii) If $\{z_i\}(t^-) \leq 0$ for all $1 \leq i \leq j$ and $\{z_{j+1}\}(t^-) > 0$, then the function $\{z_l\}(.)$ $(j+1 \leq l \leq r)$ is continuous at t.*

Proof: The values of the cones in the inclusions in (36) have been computed in the proof of Lemma 1. The result follows from Propositions 5, 6 and (36).
i) If $\{z_1\}(t^-) > 0$ then $T_\Phi^{i-1}(\{Z_{i-1}\}(t^-)) = \mathbb{R}$ and $\partial\psi_{T_\Phi^i(\{Z_i\}(t^-))}(\{z_{i+1}\}(t^+)) = \{0\}$ for all $1 \leq i \leq r$. ii) If $\{z_1\}(t^-) \leq 0$ we obtain from Proposition 5 that $\{z_1\}(t^+) = 0$. One has $T_\Phi(\{Z_1\}(t^-)) = \mathbb{R}^+$ and $\{z_2\}(t^-) > 0$. Therefore from (60) we get $\{z_2\}(t^+) = \{z_2\}(t^-) > 0$. Now $T_\Phi^2(\{Z_2\}(t^-)) = T_{T_\Phi(\{Z_1\}(t^-))}(\{z_2\}(t^-)) = T_{\mathbb{R}^+}(\{z_2\}(t^-)) = \mathbb{R}$ so $\{z_3\}(t^+) = \{z_3\}(t^-)$. iii) And so on. ■

Proposition 9 *Suppose that $CA^{r-1}B > 0$. i) If $\bar{z}_0 \succeq 0$ then $\{\bar{z}\}(0^+) = \{\bar{z}\}(0^-) = \bar{z}_0$. ii) If $t > 0$ and $\{\bar{z}\}(t^-) \succeq 0$ then $\{\bar{z}\}(t^+) = \{\bar{z}\}(t^-)$.*

Proof: i) Let $1 \leq j \leq r$ such that $\{z_1\}(0^-) = ... = \{z_{j-1}\}(0^-) = 0$ and $\{z_j\}(0^-) > 0$. From Proposition 8, we obtain that $\{z_l\}(0^+) = \{z_l\}(0^-)$, for all $j \leq l \leq r$. Let $0 \leq i \leq j-1$. We have

$$-\{z_i\}(0^-) = 0 \in \partial\psi_{T_\Phi^{i-1}(\{Z_{i-1}\}(0^-))}(\{z_i\}(0^-))$$

and

$$-\{z_i\}(0^+) = -\{z_i\}(0^+) + \{z_i\}(0^-) \in \partial\psi_{T_\Phi^{i-1}(\{Z_{i-1}\}(0^-))}(\{z_i\}(0^+)).$$

Using the monotonicity of $\partial\psi_{T_\Phi^{i-1}(\{Z_{i-1}\}(0^-))}(\{z_i\}(.))$, we get

$$\langle -\{z_i\}(0^-) + \{z_i\}(0^+), \{z_i\}(0^-) - \{z_i\}(0^+)\rangle \geq 0$$

from which we deduce that $\{z_i\}(0^+) = \{z_i\}(0^-)$. Thus $\{\bar{z}\}(0) = \{\bar{z}\}(0^-) = \{\bar{z}\}(0^+)$.

The same result holds if $\{z_1\}(0^-) = ... = \{z_r\}(0^-) = 0$. As above, we see that $\{z_i\}(0^+) = \{z_i\}(0^-)$, for all $1 \leq i \leq r-1$. Moreover, we have

$$-(CA^{r-1}B)^{-1}\{z_r\}(0^+) \in \ \partial\psi_{T_\Phi^{r-1}(\{Z_{r-1}\}(t^-))}(\{z_r\}(0^+)).$$

and also

$$-(CA^{r-1}B)^{-1}\{z_r\}(0^-) = 0 \in \ \partial\psi_{T_\Phi^{r-1}(\{Z_{r-1}\}(t^-))}(\{z_r\}(0^-))$$

so that

$$\langle (CA^{r-1}B)^{-1}\Big(-\{z_r\}(0^-) + \{z_r\}(0^+)\Big), \{z_r\}(0^-) - \{z_r\}(0^+)\rangle \geq 0$$

and thus $\{z_r\}(0^-) = \{z_r\}(0^+)$ since $(CA^{r-1}B)^{-1} > 0$.

ii) The proof of this result is similar to the one of part i). ■

Definition 2 *Let $0 \leq a < b \in \mathbb{R} \cup \{+\infty\}$ be given. We say that a solution $z \in (\mathcal{T}_{r-1}(\mathbb{R}^+))^n$ of (38)(39)(40)(43) is regular on $[a,b)$ if for each $t \in [a,b)$, there exists a right neighborhood $[t,\sigma)$ $(\sigma > 0)$ such that the restriction of $\{z\}$ to $[t,\sigma)$ is analytic.*

Lemma 3 *Suppose that $CA^{r-1}B > 0$ and let $\bar{z}_0 \succeq 0, \bar{z}_0 \neq 0$ be given. If a solution z of (38) (39) (40) (43) exists then there exists $\eta > 0$ such that $z \equiv \{z\}$ is analytic on $[0,\eta)$ and $z_1(t) > 0, \forall t \in (0,\eta)$.*

Proof: Suppose that (38)(39)(40)(43) has a solution $z = (\bar{z},\xi)^T$. Using Proposition 9, we first remark that $\{\bar{z}\}(0) = \bar{z}_0$ and since $\bar{z}_0 \succeq 0, \bar{z}_0 \neq 0$, there exists $\alpha > 0$ $(1 \leq \alpha \leq r)$ such that $\{z_\alpha\}(0) > 0$ and $\{z_k\}(0) = 0$ for all $1 \leq k \leq \alpha - 1$. We claim that there exists $\eta_0 > 0$ such that $\{z_\alpha\}$ is continuous on $[0,\eta_0]$. Indeed, suppose on the contrary that for each $\eta_0 > 0$ there exists a point $\bar{t} \in [0,\eta_0]$ such that $\{z_\alpha\}(\bar{t}^+) \neq \{z_\alpha\}(\bar{t}^-)$. Then we may find a sequence of points $\{t_i\}_{i\in\mathbb{N}}$ such that $t_i \to 0^+$ and $\{z_\alpha\}(t_i^+) \neq \{z_\alpha\}(t_i^-)$. Then using (51) and (52), we deduce that necessarily $\{z_\alpha\}(t_i) = \{z_\alpha\}(t_i^+) \leq 0, \forall i \in \mathbb{N}$. The function $\{z_\alpha\}$ is right-continuous and thus $\{z_\alpha\}(0) = \lim_{i\to\infty}\{z_\alpha\}(t_i) \leq 0$. This is a contradiction since $\{z_\alpha\}(0) > 0$. Thus $\{z_\alpha\} = z_\alpha$ is continuous on $[0,\eta_0]$. From the chain of integrators in (26) we see also that the functions $\{z_k\} = z_k \ \ (1 \le k \le \alpha - 1)$ are continuous on $(0, \eta_0)$. Therefore there exists $\eta \in (0, \eta_0)$ such that

$$z_\alpha(t) > 0, \forall t \in [0, \eta).$$

Moreover, for any $1 \le k \le \alpha - 1$, we have also

$$z_k(t) > 0, \forall t \in (0, \eta)$$

since

$$z_k(t) = z_k(0) + \int_0^t z_{k+1}(\tau)d\tau = \int_0^t z_{k+1}(\tau)d\tau.$$

In particular, $z_1(t) > 0, \forall t \in (0, \eta)$ and thus $N_{\mathbb{R}^+}(z_1(t)) = \{0\}$, $\forall t \in (0, \eta)$ and $T^i_\Phi(Z_i(t)) = \mathbb{R}$ for all $1 \le i \le r$ and all $t \in (0, \eta)$. Using $(CA^{r-1}B)^{-1} \neq 0$, we see that on $[0, \eta)$, $z = \{z\}$ is continuous and is a solution of the ODE:

$$\begin{cases} \dot{z}_1 = z_2 \\ \dot{z}_j = z_{j+1} \ \ (1 \le j \le r-1) \\ \dot{z}_r = CA^r W^{-1} z \\ \dot{\xi} = A_\xi \xi + B_\xi z_1 \\ z(0) = z_0 \end{cases} \tag{71}$$

Thus

$$z(t) = We^{At}W^{-1}z_0, \forall t \in [0, \eta)$$

and the result follows. ∎

Theorem 2 *Suppose that* $CA^{r-1}B > 0$. *If* $\bar{z}_0 \succeq 0, \bar{z}_0 \neq 0$, *then:*
i) *(Local existence) There exists* $T > 0$ *such that the system in (38) (39) (40) (43) has on* $[0, T)$ *at least one regular solution* $z \equiv \{z\}$ *given by*

$$z(t) = We^{At}W^{-1}z_0, \forall t \in [0, T);$$

ii) $z_1(t) > 0, \ \ \forall t \in (0, T)$*;*

iii) $\|z(t)\| \le e^{\|WAW^{-1}\|t}\|z_0\|, \ \ \forall\, t \in [0, T)$*;*

iv) *(Local uniqueness) If* (T^1, z^1) *and* (T^2, z^2) *are two solutions of (38) (39) (40) (43) then there exists* $T \in (0, \min\{T^1, T^2\}]$ *such that* $z^1 \equiv \{z^1\}$ *on* $[0, T)$, $z^2 \equiv \{z^2\}$ *on* $[0, T)$ *and* $\{z^1\}_{|[0,T)} = \{z^2\}_{|[0,T)}$.

Proof: Let us denote by $z = (\bar{z}, \xi)^T$ the unique solution of the ODE in (71) on $[0, +\infty)$ given by

$$z(t) = We^{At}W^{-1}z_0, \forall t \geq 0.$$

Here z is analytic on $\mathbb{R}^+$. Let α be such that $\bar{z}_{0,\alpha} > 0$ and $\bar{z}_{0,k} = \{z_k\}(0) = 0$ for all $1 \leq k \leq \alpha - 1$. We claim that there exists $\varepsilon > 0$ such that

$$z_1(t) > 0, \forall t \in (0, \varepsilon). \tag{72}$$

If $\alpha = 1$ then the result is clear. If $\alpha \geq 2$ then there exists $\varepsilon > 0$ such that $z_\alpha(t) > 0, \forall t \in (0, \varepsilon)$ and by integrating the chain of integrators in the second relations in (71), we obtain finally that (72) holds.

We may now set $T := \sup\{\tau > 0 : z_1(\tau) > 0, \forall\ \tau \in (0,t)\}$. Condition (72) entails that $N_{\mathbb{R}^+}(z_1(t)) = \{0\}$ and $T^i_\Phi(Z_i(t)) = \mathbb{R}$ for all $1 \leq i \leq r$ and for all $t \in (0, T)$. It results that z satisfies the relations in (38)-(40) (43) on $[0, T)$. Thus parts **i)** and **ii)** are proved.

This solution satisfies the ODE in (71) and thus from Gronwall's lemma (see e.g. [9]), we get

$$\|z(t)\| \leq e^{\|WAW^{-1}\|t}\|z_0\|, \quad \forall\ t \in [0, T).$$

It follows that **iii)** holds.

Let us now denote by $z = (\bar{z}, \xi)^T$ any solution of $(38) - (40), (43)$. From the proof of Lemma 3, we know that there exists $\eta > 0$ such that $z_1 = \{z_1\}$ is continuous on $[0, \eta)$ and $z_1(t) > 0, \forall t \in (0, \eta)$. Let us now set $T := \sup\{\tau > 0 : z_1(\tau) > 0, \forall\ \tau \in (0,t)\}$. We know that $T \geq \eta > 0$. So, we have $N_{\mathbb{R}^+}(z_1(t)) = \{0\}$ and $T^i_\Phi(Z_i(t)) = \mathbb{R}$ for all $1 \leq i \leq r$ and all $t \in (0, T)$. Thus $z = \{z\}$ is solution of the ODE (71) on $[0, T)$. The solution of this last problem is unique and it follows that if z is a solution of (38)-(40) that satisfies the initial condition $z(0) = z_0, \bar{z}_0 \succeq 0, \bar{z}_0 \neq 0$ then it is uniquely defined on $[0, T)$. The result in **iv)** follows. ■

Let us now discuss the case $\bar{z}_0 = 0$.

Theorem 3 *Assume that $CA^{r-1}B > 0$. If $\bar{z}_0 = 0$, then:*

i) *(Local existence) There exists $T > 0$ such that the system in (38) (39) (40) (43) has on $[0, T)$ at least one regular solution $z \equiv \{z\}$.*

ii) $\|z(t)\| \leq e^{\|WAW^{-1}\|t}\|z_0\|, \quad \forall\ t \in [0, T)$

iii) *If (T^1, z^1) and (T^2, z^2) are two regular solutions of (38) (39) (40) (43) then there exists $T \in (0, \min\{T^1, T^2\}]$ such that $z^1 \equiv \{z^1\}$ on $[0, T)$, $z^2 \equiv \{z^2\}$ on $[0, T)$ and $\{z^1\}_{|[0,T)} = \{z^2\}_{|[0,T)}$.*

Proof: Either **(a)** $d_\xi^T A_\xi^k \xi_0 = 0, \forall k \in \mathbb{N}$ or **(b)** there exists $\alpha \in \mathbb{N}$ such that $d_\xi^T A_\xi^k \xi_0 = 0, \forall\ 0 \leq k \leq \alpha - 1$ and $d_\xi^T A_\xi^\alpha \xi_0 \neq 0$.

Case **(a)**. Here we check that

$$z \equiv (\bar{z}, \xi)^T = (0, e^{A_\xi t}\xi_0)^T$$

is a regular solution of (38) (39) (40) (43). Indeed, the mapping $t \mapsto d_\xi^T e^{A_\xi t}\xi_0$ is analytic and we get here:

$$d_\xi^T \xi(t) = \sum_{k=0}^{\infty} d_\xi^T \xi^{(k)}(0) t^k = \sum_{k=0}^{\infty} d_\xi^T A_\xi^k \xi_0 t^k = 0.$$

We have thus

$$\dot{z}_i(t) - z_{i+1}(t) = 0, \forall t > 0,\ 1 \le i \le r-1$$
$$(CA^{r-1}B)^{-1}(\dot{z}_r(t) - d^T \bar{z}(t) - d_\xi^T \xi(t)) = -(CA^{r-1}B)^{-1} d_\xi^T \xi(t) = 0, \forall t > 0,$$

and

$$\dot{\xi}(t) = A_\xi \xi(t)\ (= A_\xi \xi(t) + B_\xi 0), \forall t > 0.$$

Case **(b)**. Let us first discuss the case **(b-1)** $d_\xi^T A_\xi^k \xi_0 = 0, \forall\, 0 \le k \le \alpha - 1$ and $d_\xi^T A_\xi^\alpha \xi_0 < 0$.

We check that

$$z \equiv (\bar{z}, \xi)^T = (0, e^{A_\xi t}\xi_0)^T$$

is a local regular solution of (38) (39) (40) (43). There exists $\sigma > 0$ such that

$$d_\xi^T \xi^{(\alpha)}(s) < 0, \forall s \in [0, \sigma)$$

since $d_\xi^T \xi^{(\alpha)}(0) = d_\xi^T A_\xi^\alpha \xi_0 < 0$ and $d_\xi^T \xi(.)$ is continuous.
Let $t \in (0, \sigma)$ be given. We have

$$d_\xi^T \xi^{(\alpha-1)}(t) = d_\xi^T \xi^{(\alpha-1)}(0) + \int_0^t d_\xi^T \xi^{(\alpha)}(s) ds = \int_0^t d_\xi^T \xi^{(\alpha)}(s) ds < 0$$

and so on, we get finally that

$$d_\xi^T \xi(t) < 0, \forall t \in (0, \sigma).$$

We have

$$\dot{z}_i(t) - z_{i+1}(t) = 0, \forall t \in (0, \sigma)\ \ (1 \le i \le r-1)$$
$$(CA^{r-1}B)^{-1}(\dot{z}_r(t) - d^T \bar{z}(t) - d_\xi^T \xi(t)) = -(CA^{r-1}B)^{-1} d_\xi^T \xi(t) > 0, \forall t \in (0, \sigma)$$

and thus

$$(CA^{r-1}B)^{-1}(\dot{z}_r(t) - d^T \bar{z}(t) - d_\xi^T \xi(t)) \in -\partial \psi_{T_\Phi^i(0,\dots,0)}(0) = \mathbb{R}^+,\ \forall t \in (0, \sigma).$$

Moreover

$$\dot{\xi}(t) = A_\xi \xi(t)\ (= A_\xi \xi(t) + B_\xi 0),\ \forall t \in (0, \sigma).$$

Let us now discuss the case **(b-2)** $d_\xi^T A_\xi^k \xi_0 = 0, \forall\, 0 \le k \le \alpha - 1$ and $d_\xi^T A_\xi^\alpha \xi_0 > 0$.

we check that

$$z \equiv (\bar{z}, \xi)^T = We^{At}W^{-1}z_0$$

is a local regular solution of (38) (39) (40) (43). We know that z is the solution of the system in (71). We have $z_1(0) = ... = z_r(0) = 0$ and

$$\dot{z}_r(0) = d^T \bar{z}(0) + d_\xi^T \xi(0) = d_\xi^T \xi_0 = 0.$$

Then

$$z_r^{(2)}(0) = d^T \dot{\bar{z}}(0) + d_\xi^T (A_\xi \xi_0 + B_\xi z_1(0)) =$$

$$= d_1 z_2(0) + ... + d_{r-1} z_r(0) + d_r \dot{z}_r(0) + d_\xi^T (A_\xi \xi_0 + B_\xi z_1(0)) = 0$$

and so on until

$$z_r^{(\alpha+1)}(0) = d^T \bar{z}^{(\alpha)}(0) + d_\xi^T (A_\xi^\alpha \xi_0 + B_\xi z_1^{(\alpha-2)}(0)) = d_\xi^T A_\xi^\alpha \xi_0 > 0.$$

It results that there exists $\eta > 0$ such that

$$z_r^{(\alpha+1)}(t) > 0, \forall t \in [0, \eta).$$

Then we get

$$z_i(t) > 0, \forall t \in (0, \eta) \quad (1 \leq i \leq r).$$

We may now conclude that z is a solution of (38) (39) (40) (43) on $[0, \eta)$. Indeed, if $t \in (0, \eta)$ then $N_{\mathbb{R}^+}(z_1(t)) = \{0\}$ and $-\partial\psi_{T_\Phi^{i-1}(\{Z_{i-1}\}(t^-))}(\{z_i\}(t^+)) = \{0\}$ $(2 \leq i \leq r)$, and $CA^{r-1}B \neq 0$.

The existence of a local regular solution on some interval $[0, T)$ follows.

Let us now denote by z any regular solution of (38) (39) (40) (43). We claim that there exists $T > 0$ such that z coincides on $[0, T)$ with the unique solution (see [26], [2, theorem 1] or Kato's theorem in [16]) of the system:

$$z_r(t) \geq 0, \ \forall t \geq 0, \tag{73}$$

and

$$\begin{cases} \langle \dot{z}_1(t) - z_2(t), v - z_1(t) \rangle \geq 0, \ \ \forall\, v \in \mathbb{R}, \text{ a.e. } \ t \in (0, T), \\ \langle \dot{z}_j(t) - z_{j+1}(t), v - z_j(t) \rangle \geq 0, \ \ \forall\, v \in \mathbb{R}, \text{ a.e. } \ t \in (0, T), \ \ (2 \leq j \leq r-1), \\ \langle (CA^{r-1}B)^{-1}[\dot{z}_r(t) - CA^r W^{-1} z(t)], v - z_r(t) \rangle \geq 0, \ \ \forall\, v \in \mathbb{R}^+, \ \text{ a.e. } \ t \in (0, T), \\ \langle \dot{\xi}(t) - A_\xi \xi(t) - B_\xi z_1(t), v - \xi(t) \rangle \geq 0, \ \ \forall\, v \in \mathbb{R}^{n-r}, \text{ a.e. } \ t \in (0, T), \\ z(0) = z_0. \end{cases} \tag{74}$$

We know that $z_1(0) = ... = z_r(0) = 0$ and there exists $\eta > 0$ such that the mapping $t \mapsto z_r(t)$ is analytic on $[0, \eta)$.

Either **(a)** $z_r^{(k)}(0) = 0, \forall k \in \mathbb{N}$ or **(b)** there exists $\alpha \in \mathbb{N}$ such that $z_r^{(k)}(0) = 0, \forall\, 0 \leq k \leq \alpha - 1$ and $z_r^{(\alpha)}(0) \neq 0$.

In case **(a)**, we get $z_r(t) = 0, \forall t \in [0, \eta)$ and thus $\bar{z}(t) = 0, \forall\, t \in [0, \eta)$. Then $-(CA^{r-1}B)^{-1} d_\xi \xi(t) \geq 0, \forall t \geq 0$ and $\xi(t) = e^{A_\xi t}\xi_0, \forall\, t \geq 0$ since $z_1 \equiv 0$. It results that $z(\cdot)$ is uniquely defined as the unique solution of (74) on $[0, \eta)$.

In case **(b)**, $z_r^{(\alpha)}(0) \neq 0$. It is clear that $z_r^{(\alpha)}(0) > 0$. Indeed, if we suppose on the contrary that $z_r^{(\alpha)}(0) < 0$ then there exists $\delta > 0$ such that $z_r^{(\alpha)}(t) < 0, \forall t \in [0, \delta)$. Then integrating the chain of integrators in (38) we obtain finally that $z_1(t) < 0, \forall t \in (0, \delta)$ and a contradiction. Thus $z_r^{(\alpha)}(0) > 0$ and there exists $\sigma > 0$ such that $z_r^{(\alpha)}(t) > 0, \forall t \in [0, \sigma)$. Then we obtain $z_i(t) > 0, \forall t \in (0, \sigma)\ \ (1 \leq i \leq r)$,

$$N_{\mathbb{R}^+}(z_1(t)) = \{0\}, \forall t \in (0, \sigma),$$

$$T_\Phi^i(Z_i(t)) = \mathbb{R},\ \forall t \in (0, \sigma),\ \ (1 \leq i \leq r)$$

and

$$N_{\mathbb{R}^+}(z_r(t)) = \{0\}, \forall t \in (0, \sigma).$$

This implies that here also $z(\cdot)$ is uniquely defined as the unique solution of (74) on $[0, \eta)$. Moreover, using (74) with $v = 0$ and recalling that $(CA^{r-1}B) > 0$, we see that $\frac{1}{2}\frac{d}{dt}\|z(t)\|^2 \leq WAW^{-1}z(t)$, a.e. $t \in (0, T)$. Then using Gronwall's lemma (see e.g. [9]), we obtain that $\|z(t)\| \leq e^{\|WAW^{-1}\|t}\|z_0\|,\ \forall t \in [0, T)$. ■

If $\bar{z}_0 \not\succeq 0$ then we proceed with an initial state reinitialization in requiring that $z(0^+) = z_0'$ where z_0' is uniquely defined by

$$z_{0,i}' = \text{prox}\left[T_\Phi^{i-1}(Z_{0,i-1}); z_{0,i}\right], \forall\, 1 \leq i \leq r-1,$$

$$z_{0,r}' = \text{prox}_{(CA^{r-1}B)^{-1}}\left[T_\Phi^{r-1}(Z_{0,r-1}); z_{0,r}\right]$$

and

$$z_{0,l}' = z_{0,l},\ \ (r+1 \leq l \leq n).$$

Then $\bar{z}_0' \succeq 0$ and we may apply Theorems 2 and 3 to get the following local existence result.

Corollary 1 (Local Existence and Uniqueness of a Regular Solution) *Suppose that $CA^{r-1}B > 0$. For each $z_0 \in \mathbb{R}^n$ there exists $T > 0$ such that the system in (38) (39) (40) (43) has on $[0, T)$ at least one regular solution.*

Moreover:

i) $z \equiv \{z\}$ *on* $(0, T)$

ii) $z_1 \equiv \{z_1\} \geq 0$ *on* $[0, T)$

iii) $z_i = \{z_i\} + \sum_{j=1}^{i-1}(z'_{0,j} - z_{0,j})\delta_0^{(i-j-1)}$ $(2 \le i \le r)$

iv) $\{\bar{z}\}(0^+) = \bar{z}'_0$

v) $\|\{z\}(t)\| \le e^{\|WAW^{-1}\|t}\|z'_0\|, \ \forall\, t \in [0, T)$

vi) *If* (T^1, z^1) *and* (T^2, z^2) *are two regular solutions of (38) (39) (40) (43) then there exists* $T \in (0, \min\{T^1, T^2\}]$ *such that* $\langle z^1, \varphi\rangle = \langle z^2, \varphi\rangle, \ \forall \varphi \in C_0^\infty(]0, T[; \mathbb{R}^n)$.

■

Let us now provide a global existence and uniqueness result.

Corollary 2 (Global Existence and Uniqueness of a Regular Solution) *Suppose that* $CA^{r-1}B > 0$. *For each* $z_0 \in \mathbb{R}^n$, *the system in (38) (39) (40) (43) has at least one regular solution.*

Moreover:

i) $z_1 \equiv \{z_1\} \ge 0$ *on* $\mathbb{R}^+$

ii) $\{\bar{z}\}(0^+) = \bar{z}'_0$

iii) $\|\{z\}(t)\| \le e^{\|WAW^{-1}\|t}\|z_0\|, \ \forall\, t \in \mathbb{R}^+$

iv) *If* z^1 *and* z^2 *are two regular solutions of (38) (39) (40) (43) then* $\langle z^1, \varphi\rangle = \langle z^2, \varphi\rangle, \ \forall \varphi \in C_0^\infty(\mathbb{R}^+; \mathbb{R}^n)$.

Proof: There exists $T > 0$ such that the system in (38)(39)(40)(43) has on $[0, T)$ a unique regular solution z that satisfies the relations **(i)-(vi)** of Corollary 1. From **(i)** and **(v)**, we know that $\|z(t)\| \le e^{\|WAW^{-1}\|t}\|z'_0\|, \forall t \in (0, T)$. Thus $\|z(T^-)\| \le e^{\|WAW^{-1}\|t}\|z'_0\| < +\infty$ and we may then proceed with a state reinitialization by requiring that $z(T^+) = z'_1$, where z'_1 is uniquely defined by

$$z'_{1,i} = \text{prox}\left[T_\Phi^{i-1}(Z_{i-1}(T^-)); z_i(T^-)\right], \ (1 \le i \le r-1),$$

$$z'_{1,r} = \text{prox}_{(CA^{r-1}B)^{-1}}\left[T_\Phi^{r-1}(Z_{1,r-1}(T^-)); z_r(T^-)\right],$$

$$z'_{1,l} = z_l(T^-), \ \ (r+1 \le l \le n).$$

From Corollary 1, we get a prolongation of $z(\cdot)$ as a unique regular solution of the system in (38)(39)(40)(43) on $[0, T + T_1)$ with $T_1 > 0$. We have

$$\|\{z(t)\}\| \le e^{\|WAW^{-1}\|t}\|z'_1\|, \forall t \in [T, T_1).$$

We also have

$$\mid z'_{1,i} \mid \le |\text{prox}\left[T_\Phi^{i-1}(Z_{i-1}(T^-)); z_i(T^-)\right]| \le \ \mid z_i(T^-) \mid, \ \ (1 \le i \le r-1),$$

$$\langle (CA^{r-1}B)^{-1}z'_{1,r}, z'_{1,r}\rangle \leq \langle (CA^{r-1}B)^{-1}z_r(T^-), z_r(T^-)\rangle$$

and thus

$$\mid z'_{1,r} \mid \leq \mid z_r(T^-) \mid,$$

since $(CA^{r-1}B)^{-1} > 0$. Moreover,

$$\mid z'_{1,l} \mid = \mid z_l(T^-) \mid, \quad (r+1 \leq l \leq n).$$

Consequently $\|z'_1\| \leq \|z(T^-)\|$ and

$$\|\{z(t)\}\| \leq e^{\|WAW^{-1}\|(t-T)}\|z'_1\| \leq e^{\|WAW^{-1}\|(t-T)}\|z(T^-)\| \leq$$

$$\leq (e^{\|WAW^{-1}\|(t-T)})(e^{\|WAW^{-1}\|T})\|z'_0\| = e^{\|WAW^{-1}\|t}\|z'_0\|, \forall t \in [T, T_1).$$

We have also

$z_i = \{z_i\} + \sum_{j=1}^{i-1}(z'_{0,j} - z_{0,j})\delta_0^{(i-j-1)} + \sum_{j=1}^{i-1}(z'_{1,j} - z_j(T^-))\delta_0^{(i-j-1)} \quad (2 \leq i \leq r)$

$\{\bar{z}\}(T^+) = \bar{z}'_1$,

$z \equiv \{z\}$ on $(0,T) \cup (T, T_1)$,

$z_1 = \{z_1\} \geq 0$ on $[0, T_1)$.

As above, we check that $\|z'_0\| \leq \|z_0\|$, so that

$$\|\{z(t)\}\| \leq e^{\|WAW^{-1}\|t}\|z_0\|, \forall t \in [0, T_1).$$

And so on, this procedure yields a unique maximal regular solution z on $[0, T_{max})$ such that $\|z(t)\| \leq e^{\|WAW^{-1}\|t}\|z'_0\|$, $\forall t \in (0, T_{max})$. Here $T_{max} = +\infty$ since otherwise, we would obtain $|z(T_{max})| \leq e^{\|WAW^{-1}\|T_{max}}\|z_0\| < +\infty$ and Corollary 1 furnishes a prolongation of z as a solution of (38)(39)(40)(43), which contradicts the definition of T_{max}.

Applying corollary 2 and proposition 7 it follows that autonomous "dissipative" higher order sweeping processes are well-posed systems. An example is detailed in section 6.

Remark 13 i) *Let $0 = T_0 < T_1 < T_2 < \ldots$ be the sequence of non-trivial ($\{z\}(T_i^-) \neq \{z\}(T_i^+)$) state reinitialization points required to get the global solution z. Then $z_1 = \{z_1\}$ and*

$$z_i = \{z_i\} + \sum_{\alpha}\sum_{j=1}^{i-1}(\{z_j\}(T_\alpha^+) - \{z_j\}(T_\alpha^-))\delta_{T_\alpha}^{(i-j-1)} \quad (2 \leq i \leq r)$$

ii) *Corollary 1 and Corollary 2 can be generalized to the case $m \geq 2$ and uniform relative degree provided that one supposes that $CA^{r-1}B$ is a Stieltjes matrix, i.e. a nonsingular symmetric M-matrix [8]. These assumptions ensure that the matrix $CA^{r-1}B$ is positive definite and $(CA^{r-1}B)^{-1}$ is nonnegative in the sense that $(CA^{r-1}B)_{ij}^{-1} \geq 0$ for all $i, j \in$*

$\{1,...,n\}$. *Note that if* $m \geq 2$*, for* $z = (z_1,...z_r,\xi)^T \in \mathbb{R}^n$*,* $(z_i \in \mathbb{R}^m$*,* $\xi \in \mathbb{R}^{n-rm})$*, it would be convenient to use the norm*

$$\|z\| := \sqrt{\sum_{\alpha=1}^{r-1} \|z_\alpha\|_m^2 + \langle CA^{r-1}Bz_r, z_r\rangle + \|\xi\|_{n-mr}^2}.$$

iii) *Let us notice that the result given by Corollary 2 does not encompass the bouncing ball of Mechanics. However mechanical systems of the form*

$$\begin{aligned} &M\ddot{q}(t) + F\dot{q}(t) + Kq(t) = L^T\lambda \\ &0 \leq \lambda(t) \perp Lq(t) \geq 0 \end{aligned} \tag{75}$$

can be treated in our framework, assuming that $q(0) \geq 0$ *and that* $FM^{-1}F^T (= CAB)$ *has the required properties.*

4.6 Sweeping process vs. linear complementarity systems

LCSs consist of (16) with a complementarity relation $0 \leq w(t) \perp \lambda(t) \geq 0$ [30, 28]. It is clear that the formalism proposed in this paper involves complementarity conditions between some variables, see Theorem 1 and Proposition 4. However the work in [30, 28] and the work in this paper differ in several fundamental ways. First a local well-posedness result has been obtained in [30, Theorem 6.3] for $m \geq 1$, and a global well-posedness result has been obtained in [28, Theorem 3.6.10] for the case $m = 1$. Therefore, from item **ii)** in remark 13, our results are more general. Nevertheless the major discrepancy is that the approach in [30, 28] relies on an event-driven point of view on hybrid systems, as is illustrated in their definition of a solution [30, Definition 4.10]. In other words, the system is initialized in a so-called mode (a DAE), and then integration progresses until the system has to switch to another mode. Especially, this way of thinking, leads to event-driven time-integration schemes. No numerical algorithm is proposed in [30] but only a rough guideline for constructing such a scheme [30, §7]. It is well-known that event-driven schemes possess strong drawbacks such as the impossibility to pass through finite accumulations of events, and the impossibility to obtain convergence proofs [15]. The point of view that is adopted here, is dramatically different. It is the point of view of (unbounded) differential inclusions, which naturally yields time-stepping numerical schemes and paves the way towards convergence proofs, as illustrated in section 5.

To end this section, let us notice that if the complementarity relation is of the form $0 \leq w(t) = Cx(t) + D\lambda(t) \perp \lambda(t) \geq 0$, and if the matrix D is full rank, then the system has relative degree 0. In such a case the complementarity relation makes perfect sense, since $\lambda(t)$ is a function. Assume for instance that $D = D^T > 0$. Then it is deduced from convex analysis (see e.g. [62, Example 10.2]) that λ is the projection of $-D^{-1}Cx$ on $(\mathbb{R}^+)^m$ in the metric defined by D. As such, it is a Lipschitz continuous function of x [25]. Thus the LCS is an ODE. This clearly shows why the framework of differential inclusions is not well suited to the case $r = 0$.

5 A solution method for the numerical time integration: The EMTS scheme

Following the work of Moreau [48, 49, 51, 54, 56] and co-workers [45, 38, 32], we aim at designing the so-called "Extended Moreau's Time-Stepping" (EMTS) scheme. This time-stepping scheme i.e., without an explicit procedure for handling the times of events is based on the approximation on a time interval of the Measure Differential Formalism of the Higher Order Moreau's Sweeping Process (44)–(46) (43).

This section is organized as follows. In Section 5.1, we recall basic facts about the numerical time-integration of the Moreau's sweeping process with a relative degree less or equal to 2. In Section 5.2, the principle of the construction of an approximated solution is stated. In the spirit of the work in [45], several major properties of the proposed scheme are outlined in Section 5.3 that pave the way to a convergence result. In Section 5.4, we give an overview of a possible implementation of the numerical scheme. Finally, we conclude by giving some numerical applications in Section 5.5. Particularly, the numerical scheme is compared to a direct use of a Backward Euler scheme as in [17] (Section 5.5.1). Finally, the influence of the zero-dynamics is highlighted (Section 5.5.2) and the empirical order of the scheme is evaluated (Section 5.5.3) on a particular example.

The following notation is used throughout this section. We denote by $0 = t_0 < t_1 < \ldots < t_k < t_N = T$ a finite partition (or a subdivision) of the time interval $[0, T], T > 0$. The integer N stands for the number of time intervals in the subdivision. The length of a time step is denoted by $h_k = t_{k+1} - t_k$. For simplicity sake, we consider only in the sequel a constant time length h. The approximation of $f(t_k)$, the value of a real function f at the time t_k, is denoted by f_k .

5.1 Background on the numerical time integration of the Moreau's sweeping process

In this section, we give some details about the seminal work of Moreau on the numerical time integration of the sweeping process.

5.1.1 First order sweeping process

Let us consider the first order sweeping process, or more precisely, the sweeping process of relative degree 1 introduced in (1) :

$$-dz \in N_{K(t)}(z(t)) \;\; (t \geq 0), \;\; z(0) \in \mathbb{R}^n, \tag{76}$$

Under suitable hypothesis on the multivalued function $t \mapsto K(t)$, numerous convergence and consistency results [45, 38] have been given together with well-posedness results using the so-called "Catching-up algorithm" defined in [51]:

$$-\left(z(t^+_{k+1}) - z(t^+_k)\right) \in \partial\psi_{K(t^+_{k+1})}(z(t^+_{k+1})) \tag{77}$$

By elementary convex analysis, and using the convention that $z_{k+1} = z(t_{k+1}^+)$ this is equivalent to:

$$z_{k+1} = prox(K_{k+1}, z_k) \tag{78}$$

Contrary to the standard backward Euler scheme with which it might be confused, the catching-up algorithm is based on the evaluation of the measure dz on the interval $]t_k, t_{k+1}]$, i.e. $dz(]t_k, t_{k+1}]) = z(t_{k+1}^+) - z(t_k^+)$. Indeed, the backward Euler scheme is based on the approximation of $\dot{z}(t)$ which is not defined in a classical sense for our case. When the time step vanishes, the approximation of the measure dz tends to a finite value corresponding to the jump of z. This remark is crucial for the consistency of the scheme. Particularly, this fact ensures that we handle only finite values.

In the same way, using higher order numerical schemes is at best useless, more often it is dangerous. Basically, a general way to obtain a finite difference-type scheme of order n is to write a Taylor expansion of order n or higher. Such a scheme is meant to approximate the n-th derivative of the discretized function. If the solution we are dealing with is obviously not differentiable, what is the meaning of using a scheme with order $n \geq 2$? Such a scheme will try to approximate derivatives which do not exist. At the times of non-differentiability, it may introduce in the solution artificial unbounded terms creating oscillations, etc. In summary, higher-order numerical schemes are inadequate for time-stepping discretization of complementarity systems.

5.1.2 Second order sweeping process: Overview of the contact dynamics method

The "Non Smooth Contact Dynamics (NSCD)" method [56, 57, 32] is the numerical discretization of the second order Moreau Sweeping process introduced by Moreau in [54, 53] in the context of Lagrangian mechanical systems (81) reformulated as a measure differential inclusion [64, 53, 54] :

$$M(q)dv + F(q, v, t)dt = \sum_{\alpha=1}^{\nu} \nabla h_\alpha(q)\lambda_\alpha, \tag{79}$$

where q is the vector of absolutely continuous generalized coordinates, dv is a differential measure associated with the velocity $v = \dot{q}^+$ considered as a RCLBV function, dt is the Lebesgue measure $M(q)$ is the mass matrix and $F(q, v, t)$ is the set of forces acting upon the system, $h_\alpha(q) \geq 0$, $\alpha = 1 \ldots \nu$, are the unilateral constraints and finally, λ is a measure. To complete this measure differential equation, Moreau proposed a compact formulation of an inelastic impact law as a measure inclusion:

$$-H^T(q(t))\lambda \in \partial\psi_{V(q(t))}(v(t^+) + ev(t^-)) \tag{80}$$

where $V(q)$ is the tangent cone to $\Phi(t)$ at q, e is the coefficient of restitution and $H^T(q)\lambda = \sum_{\alpha=1}^{\nu} \nabla h_\alpha(q)\lambda_\alpha$. Finally, we obtain a MDI, the so-called sweeping process:

$$M(q)dv + F(q, v, t)dt \in -\partial\psi_{V(q)}(v(t^+) + ev(t^-)) \tag{81}$$

The NSCD method performs the numerical time integration of the MDI (81) on a interval $(t_k, t_{k+1}]$. Using the notation

$$v_{k+1} \approx v(t^+_{k+1}); \quad \mu_{k+1} \approx \lambda(]t_k, t_{k+1}]) \tag{82}$$

it may be written down as follows:

$$\begin{cases} -M(v_{k+1} - v_k) - h((1-\theta)F(v_{k+1}, q_{k+1}, t_{k+1}) - \theta F(v_k, q_k, t_k)) = H^T(\tilde{q}_{k+1})\mu_{\alpha\, k+1} \\ H^T(\tilde{q}_{k+1})\mu_{\alpha\, k+1} \in \partial\psi_{T_\Phi(\tilde{q}_{k+1})}(v_{k+1}) \\ q_{k+1} = q_k + h((1-\theta)v_{k+1} + \theta v_k), \quad \theta \in [\frac{1}{2}, 1] \\ \tilde{q}_{k+1} = q_k + hv_k \end{cases} \tag{83}$$

The inclusion can be stated equivalently as a complementarity problem:

$$\text{if } h(\tilde{q}_{k+1}) \leq 0 \text{ then } 0 \leq H(\tilde{q}_{k+1})v_{k+1} \perp \mu_{k+1} \geq 0 \tag{84}$$

The value $\tilde{q}_{k+1}$ is a prediction of the position which allows the computation of the tangent cone T_Φ. A $\theta-$method is used for the integration of the position assuming that q is absolutely continuous. The same approximation is made with the term $F(v(t^+), q(t), t)$.

Comments From a numerical point of view, two major lessons can be learned form this work:

1. First, the various terms manipulated by the numerical algorithm are of finite value. The use of differential measures of the time interval $(t_k, t_{k+1}]$,. i.e., $dv(]t_k, t_{k+1}]) = v(t^+_{k+1}) - v(t^+_k)$ and $= \mu_{k+1} = d\lambda(]t_k, t_{k+1}])$ is fundamental and allows a rigorous treatment of the non smooth evolutions. When the time-step h vanishes, it allows to deal with finite jumps. When the evolution is smooth, the scheme is equivalent to a backward Euler scheme. We can remark that nowhere an approximation of the acceleration is used.

2. Secondly, the inclusion in terms of velocity allows us to treat the displacement as a secondary variable. A viability lemma ensures that the constraints on q will be respected at convergence. We will see further that this formulation gives more stability to the scheme.

These remarks on the contact dynamics method might be viewed only as some numerical tricks. In fact, the mathematical study of the second order MDI by Moreau provides a sound mathematical ground to this numerical scheme. It is noteworthy that convergence results have been proved for such time-stepping schemes [45, 69].

Example of the bouncing ball The NSCD method provides a numerical scheme with very nice properties. The reader may convince his/herself of this by studying the simple bouncing ball on a rigid plane subject to gravity and with elastic restitution. The proposed time-discretization of the motion of this ball is

$$-m(v(t_{k+1})+v(t_k))-hmg\in\partial\psi_{V(\tilde{q}_{k+1})}(v(t_{k+1})+ev(t_k)) \tag{85}$$

where m is the mass of the ball, e the coefficient of restitution and g the gravity. One notes that the dissipativity property shows through the power of h in the term hg which has the dimension of an impulse (there is no h pre-multiplying the right-hand-side since this is a cone).

If $q_0>0$ then the ball falls down until penetration is detected at step k^* (i.e. $q_{k^*-1}>0$ while $q_{k^*}<0$). Then the velocity is reversed, i.e. $v_{k^*+1}=-ev_{k^*}$ while $q_{k^*+1}=q_{k^*}$. Thus the system is re-initialized at each impact, with the same velocity and at the same position. There are no errors introduced by the numerical scheme and one can simulate several billions of such cycles with neither energy gain nor losses. Clearly this is not possible with an event-driven scheme, even if a very accurate detection procedure is used. The unavoidable penetration is not a major issue, since anyway the discretized system cannot be exactly at $q=0$. What is crucial is that the penetration goes to zero when $h\to 0$. In the case $e\in(0,1)$, an infinity of rebounds in finite time occurs in the continuous time model. This Zeno behaviour is correctly integrated as depicted on Figure 1.

5.2 Principle

Let us start with a generic equation of the measure differential formalism for the extended sweeping process (44) for $1\le i\le r-1$:

$$\begin{aligned} &dz_i - z_{i+1}(t)dt = d\nu_i,\\ &d\nu_i \in -\partial\psi_{T_\Phi^{i-1}(Z_{i-1}(t^-))}(z_i(t^+)). \end{aligned} \tag{86}$$

As in Section 2, it results from Proposition 1 that an evaluation of this MDI on the time interval $(t_k,t_{k+1}]$ yields :

$$\begin{aligned} &dz_i((t_k,t_{k+1}]) - \textstyle\int_{(t_k,t_{k+1}]} z_{i+1}(\tau)d\tau = d\nu_i((t_k,t_{k+1}])\\ &d\nu_i((t_k,t_{k+1}]) \in \overline{\mathrm{co}}(\cup_{\tau\in(t_k,t_{k+1}]} -\partial\psi_{T_\Phi^{i-1}(Z_{i-1}(\tau^-))}(z_i(\tau^+))). \end{aligned} \tag{87}$$

The values of the measures $dz_i((t_k,t_{k+1}])$ and $\mu_{i,k+1}\triangleq d\nu_i((t_k,t_{k+1}])$ are kept as primary variables and this fact is crucial for the consistency of the method for the nonsmooth evolutions. The integral term is approximated thanks to :

$$\int_{(t_k,t_{k+1}]} z_{i+1}(\tau)d\tau \approx hz_{i+1}(t_{k+1}^+) = hz_{i+1,k+1} \tag{88}$$

and then we obtain :

$$z_{i,k+1} - z_{i,k} - hz_{i+1,k+1} = \mu_{i,k+1}. \tag{89}$$

For the approximation of the inclusion, the union of convex cones is approximated in the following way :

$$\overline{\text{co}}(\cup_{\tau \in (t_k, t_{k+1}]} - \partial \psi_{T_\Phi^{i-1}(Z_{i-1}(\tau^-))}(z_i(\tau^+))) \approx -\partial \psi_{T_\Phi^{i-1}(Z_{i-1}(t_k^-))}(z_i(t_{k+1}^+)). \tag{90}$$

Assuming, as in (88), that the approximation of z_i is constant on each interval $(t_k, t_{k+1}]$, we get :

$$\mu_{i,k+1} \in -\partial \psi_{T_\Phi^{i-1}(Z_{i-1,k})}(z_{i,k+1}). \tag{91}$$

Finally, the time integration of a generic equation of the MDI in (44) is given by :

$$z_{i,k+1} - z_{i,k} - hz_{i+1,k+1} = \mu_{i,k+1} \in -\partial \psi_{T_\Phi^{i-1}(Z_{i-1,k})}(z_{i,k+1}) \ \ (1 \leq i \leq r-1). \tag{92}$$

The last equation (45) is discretized in the same way :

$$\begin{aligned} &z_{r,k+1} - z_{r,k} - hCA^r W^{-1} z_{k+1} = CA^{r-1}B \ \mu_{r,k+1} \\ &\mu_{r,k+1} \in -\partial \psi_{T_\Phi^{r-1}(Z_{r-1,k})}(z_{r,k+1}). \end{aligned} \tag{93}$$

For the zero dynamics defined in (46), we use for the sake of simplicity[1] an Euler Backward scheme :

$$\xi_{k+1} - \xi_k - hA_\xi \xi_{k+1} - hB_\xi z_{1,k+1} = 0. \tag{94}$$

Definition 3 (Extended Moreau's Time stepping (EMTS) scheme) *The numerical time integration of the Higher Order Sweeping Process* (44)–(45)–(46) *is defined as the inclusion in* (92)*,* (93) *and* (94)*.*

The following notation is used for the discretized variables. Let us denote the discretized state vector by

$$z_{k+1} = [z_{1,k+1}, \ldots, z_{r,k+1}, \xi_{k+1}^T]^T = [\bar{z}_{k+1}^T, \xi_{k+1}^T]^T$$

the vector of discretized multipliers $\mu_{i,k+1}, 1 \leq i \leq r$ by μ_{k+1} :

$$\mu_{k+1} = [\mu_{1,k+1}, \ldots, \mu_{r,k+1}]^T.$$

Then the discrete-time system in (92), (93) and (94) can be rewritten compactly as (see (66) and (67))

$$z_{k+1} - z_k = hWAW^{-1} z_{k+1} + \bar{G}\mu_{k+1} \tag{95}$$

(which is the discrete-time counterpart to (70)).

[1] Depending on the regularity of z_1, a higher order scheme might be used for the time-integration of the zero dynamics.

Comments As we have seen earlier, the measures of the time interval $(t_k, t_{k+1}]$, i.e. $dz((t_k, t_{k+1}])$ and $\mu_{i,k+1} \stackrel{\Delta}{=} d\nu_i((t_k, t_{k+1}])$ are kept as primary variables. This fact ensures that the various terms manipulated by the numerical algorithm are of finite values. The use of differential measures of the time interval $(t_k, t_{k+1}]$ allows a rigorous treatment of the nonsmooth evolutions. When the time-step h vanishes, it allows to deal with finite jumps. When the evolution is smooth, the scheme is equivalent to a backward Euler scheme. We can remark that nowhere a direct approximation of the density z'_t with respect to the Lebesgue measure is made. The use of a first order algorithm is not chosen as usual through the approximation of the integral term (88) but required by the evaluation of the differential measure.

As it has been observed in Remark 12, it is possible to incorporate a different jump mapping. This is also the case for the numerical time integration in approximating the mapping (62) by :

$$\mu_{i,k+1} \in -\partial\psi_{T_\Phi^{i-1}(Z_{i-1,k})}\left(\frac{z_{i,k+1} - e_{i+1}z_{i,k}}{1+e_{i+1}}\right). \tag{96}$$

In the case of Lagrangian systems, the function z_1 is assumed to be absolutely continuous and the measure $d\nu_1$ is chosen identically equal to zero. So in this case we have to take care that μ_1 vanishes when the time step vanishes also. An other way is to choose $\mu_1 \equiv 0$. In this case slight violation of the constraint , increasing with the order of the sweeping process is expected.

5.3 Properties of the discrete-time extended sweeping process

We therefore consider here the discrete-time system in (92), (93) and (94). In this section, some important properties are shown which are thought to pave the way towards a convergence proof of the discrete-time solutions towards a solution of the continuous-time sweeping process. In the case $r = 2$ and $z_1(0) \geq 0$, such a convergence has been established in [45, §3.2]. The hardest part of the proof is not convergence towards some limit itself, but showing that the limit *is* a solution of the sweeping process. One discrepancy with respect to the works in [45, §3.2] [43, 69] is that the well-posedness of the higher order sweeping process has already been proved in corollary 2.

5.3.1 A dissipation inequality

In this section we assume that the triple $(WAW^{-1}, \bar{G}, H)$ is positive real and the corresponding realisation is minimal as in Proposition 7.

Proposition 10 *Let us consider* (92)*,* (93) *and* (94)*, and the matrix J in (68). Then*

$$\frac{1}{2}z_{k+1}^T J z_{k+1} - \frac{1}{2}z_k^T J z_k \leq -\frac{1}{2}(z_{k+1} - z_k)^T J (z_{k+1} - z_k) + h z_{k+1}^T J W A W^{-1} z_{k+1} \tag{97}$$

for all $k \geq 0$.

Proof:
Let us premultiply both sides of (95) by $z_{k+1}^T J$. This gives

$$z_{k+1}^T J z_{k+1} - z_{k+1}^T J z_k = h z_{k+1}^T JWAW^{-1} z_{k+1} + z_{k+1}^T J\bar{G}\mu_{k+1}.$$

We have

$$z_{k+1}^T J z_{k+1} - z_{k+1}^T J z_k = \frac{1}{2} z_{k+1}^T J z_{k+1} - \frac{1}{2} z_k^T J z_k + \frac{1}{2}(z_{k+1} - z_k)^T J (z_{k+1} - z_k) \tag{98}$$

and

$$h z_{k+1}^T JWAW^{-1} z_{k+1} + z_{k+1}^T J\bar{G}\mu_{k+1} \leq h z_{k+1}^T JWAW^{-1} z_{k+1} \tag{99}$$

where we used the fact that $z_{k+1}^T J\bar{G}\mu_{k+1} \leq 0$ as a consequence of the relations in (91) and (93) (recall that $(J\bar{G})^T = (I_r \; 0_{(n-r)\times r})$). From (98) and (99) the result follows. ■

5.3.2 Boundedness properties

Proposition 11 *There exists a constant $\alpha > 0$ such that for all $h > 0$ and all $k \geq 1$, $||z_k|| \leq \alpha$. Moreover, for any given $h^* > 0$, there exists a constant $M \equiv M(h^*) > 0$ such that $||\bar{G}\mu_k|| \leq M, \forall h \in]0, h^*[$.*

Proof: Notice that from (97), we get

$$z_{k+1}^T J z_{k+1} \leq z_k^T J z_k + z_{k+1}^T (JWAW^{-1} + W^{-T} A^T W^T J) z_{k+1} - \frac{1}{2}(z_{k+1} - z_k)^T J (z_{k+1} - z_k) \leq z_k^T J z_k, \tag{100}$$

where we have used the facts that the matrix $JWAW^{-1} + W^{-T}A^T W^T J$ is negative semi-definite and the matrix J is positive definite . Proceeding by induction one concludes that

$$z_{k+1}^T J z_{k+1} \leq z_0^T J z_0 \;\; (\forall k \geq 0).$$

Let us denote by $\lambda_J > 0$ the smallest eigenvalue of J. We get

$$\lambda_J ||z_{k+1}||^2 \leq z_0^T J z_0 \;\; (\forall k \geq 0)$$

and the result is proved with $\alpha = \sqrt{\frac{z_0^T J z_0}{\lambda_J}}$. From (95) one can choose $M = 2\alpha + h^* ||WAW^{-1}||\alpha$. ■

5.3.3 Local bounded variation

What follows is strongly inspired from [45, Lemma 2.5]. We first notice that since all the cones $T_\Phi^i(\cdot)$ are either $I\!R$ or $I\!R^+$, it follows that the closed ball $\bar{B}(a, R) = \{z \in \mathbb{R} : \; ||\bar{z} - a|| \leq R\} \subset T_\Phi^i(\cdot)$ for any $a > 0$ and $R < \frac{a}{2}$. We define $z_i^N(\cdot)$ as the step function $[0, T] \rightarrow I\!R$ such that $z_i^N(t) = z_{i,k}$ for all $t \in [t_k, t_{k+1})$, $0 \leq k \leq N-1$, $1 \leq i \leq r$.

Proposition 12 *The total variation of $z_i^N(\cdot)$, $1 \leq n$, in $[0,T)$ is bounded above according to:*

$$\begin{aligned} var(z_i^N,[0,T]) &\leq \tfrac{1}{2R}|z_{i,0}-a|^2 + \tfrac{\alpha^2}{2R}T^2 + \alpha T(1+\tfrac{1}{R}|z_{i,0}-a|) \qquad (1 \leq i \leq r-1) \\ var(z_r^N,[0,T]) &\leq \tfrac{1}{2R}|z_{r,0}-a|^2 + \tfrac{\beta^2\alpha^2}{2R}T^2 + \beta\alpha T(1+\tfrac{1}{R}|z_{1,0}-a|) \\ var(\xi^N,[0,T]) &\leq (\gamma+\delta)\alpha T \end{aligned} \tag{101}$$

where $||CA^rW^{-1}|| \leq \beta$, $||A_\xi|| \leq \gamma$ and $||B_\xi|| \leq \delta$, whereas α is as in Proposition 11. Moreover there exists a constant $K>0$ such that for all $N \in \mathbb{N}$:

$$var(z^N,[0,T]) \leq K. \tag{102}$$

Proof: In this proof we suppress the N in z_i^N and ξ^N to simplify the notation. Recall that

$$z_{i,k+1} - z_{i,k} - hz_{i+1,k+1} \in -\partial\psi_{T_\Phi^{i-1}(Z_{i-1,k})}(z_{i,k+1}) \ \ (1 \leq i \leq r-1)$$

and

$$z_{r,k+1} - z_{r,k} - hCA^rW^{-1}z_{k+1} \in -CA^{r-1}B\partial\psi_{T_\Phi^{r-1}(Z_{r-1,k})}(z_{r,k+1}) = -\partial\psi_{T_\Phi^{r-1}(Z_{r-1,k})}(z_{r,k+1})$$

where we have used the facts that $CA^{r-1}B > 0$ and that $\partial\psi_{T_\Phi^{r-1}(Z_{r-1,k})}(z_{r,k+1})$ is a closed cone. Then we may write

$$z_{i,k+1} = prox[T_\Phi^{i-1}(Z_{i-1,k}); z_{i,k} + hz_{i+1,k+1}] \ \ (1 \leq i \leq r-1)$$

and

$$z_{r,k+1} = prox[T_\Phi^{r-1}(Z_{r-1,k}); z_{r,k} + hCA^rW^{-1}z_{k+1}].$$

Considering the ball $\bar{B}(a,r)$, using (92), (93) and [45, Lemma 0.4.4] (see appendix B) one gets

$$\begin{aligned} |z_{i,k+1}-a| &\leq |z_{i,0}-a| + h(k+1)\alpha, \qquad (1 \leq i \leq r-1, k \geq 0), \\ |z_{r,k+1}-a| &\leq |z_{r,0}-a| + h(k+1)\beta\alpha. \end{aligned} \tag{103}$$

Let us set $w_i = z_{i,0} + hz_{i+1,1}$ and $w_r = z_{r,0} + hCA^rW^{-1}z_1$. One has $|z_{i,1} - z_{i,0}| \leq |z_{i,1} - w_i| + |w_i - z_{i,0}|$ and by [45, Lemma 0.4.3] (see appendix A):

$$\begin{aligned} |z_{i,1}-z_{i,0}| &\leq \tfrac{1}{2R}(|w_i-a|^2 - |z_{i,1}-a|^2) + h\alpha, \quad (1 \leq i \leq r-1), \\ |z_{r,1}-z_{r,0}| &\leq \tfrac{1}{2R}(|w_r-a|^2 - |z_{r,1}-a|^2) + h\beta\alpha. \end{aligned} \tag{104}$$

More generally, we get the inequalities

$$|z_{i,k+1} - z_{i,k}| \leq \frac{1}{2R}(|z_{i,k} + hz_{i+1,k+1} - a|^2 - |z_{i,k+1} - a|^2) + h\alpha \leq$$

$$\leq \frac{1}{2R}(|z_{i,k} - a|^2 + h^2|z_{i+1,k+1}|^2 + 2h|z_{i+1,k+1}|\ |z_{i,k} - a| - |z_{i,k+1} - a|^2) + h\alpha$$

$$\leq \frac{1}{2R}(|z_{i,k} - a|^2 + h^2\alpha^2 + 2h\alpha\ |z_{i,k} - a| - |z_{i,k+1} - a|^2) + h\alpha$$

for $1 \leq i \leq r-1$, and

$$|z_{r,k+1} - z_{r,k}| \leq \frac{1}{2R}(|z_{r,k} + hCA^rW^{-1}z_{k+1} - a|^2 - |z_{r,k+1} - a|^2) + h\alpha\beta.$$

$$\leq \frac{1}{2R}(|z_{r,k} - a|^2 + h^2|CA^rW^{-1}z_{k+1}|^2 + 2h|CA^rW^{-1}z_{k+1}|\ |z_{r,k} - a| - |z_{r,k+1} - a|^2) + h\beta\alpha \leq$$

$$\leq \frac{1}{2R}(|z_{r,k} - a|^2 + h^2\beta^2\alpha^2 + 2h\beta\alpha\ |z_{r,k} - a| - |z_{r,k+1} - a|^2) + h\beta\alpha.$$

Now using (103) we get $|z_{i,k} - a| \leq |z_{i,0} - a| + hk\alpha$ for all $1 \leq i \leq r-1$, and $|z_{r,k} - a| \leq |z_{r,0} - a| + hk\beta\alpha$, and setting $T_k = kh$ (so that $T = Nh$), we get

$$|z_{i,k+1} - z_{i,k}| \leq \frac{1}{2R}(|z_{i,k} - a|^2 - |z_{i,k+1} - a|^2) + \alpha h \left(1 + \frac{1}{R}|z_{i,0} - a|\right) + \frac{\alpha^2}{2R}(h^2 + 2hT_k) \quad (105)$$

for $1 \leq i \leq r-1$, and

$$|z_{r,k+1} - z_{r,k}| \leq \frac{1}{2R}(|z_{r,k} - a|^2 - |z_{r,k+1} - a|^2) + \alpha\beta h \left(1 + \frac{1}{R}|z_{r,0} - a|\right) + \frac{\alpha^2\beta^2}{2R}(h^2 + 2hT_k). \quad (106)$$

We claim that

$$\sum_{k=0}^{j-1} |z_{i,k+1} - z_{i,k}| \leq \frac{1}{2R}(|z_{i,0} - a|^2 - |z_{i,j} - a|^2) + \alpha T_j \left(1 + \frac{1}{R}|z_{i,0} - a|\right) + \frac{\alpha^2}{2R}T_j^2 \quad (107)$$

for $1 \leq i \leq r-1$, and

$$\sum_{k=0}^{j-1} |z_{r,k+1} - z_{r,k}| \leq \frac{1}{2R}(|z_{r,0} - a|^2 - |z_{r,j} - a|^2) + \alpha\beta T_j \left(1 + \frac{1}{R}|z_{r,0} - a|\right) + \frac{\alpha^2\beta^2}{2R}T_j^2. \quad (108)$$

We prove the result by induction. The result holds for $j = 1$ as a direct consequence of (105) and (106) with $k = 0$. Suppose now that (107) holds at step j. We claim that (107) holds at step $j+1$. Indeed, using again (105), we get for step $j+1$:

$$\sum_{k=0}^{j} |z_{i,k+1} - z_{i,k}| = (\sum_{k=0}^{j-1} |z_{i,k+1} - z_{i,k}|) + |z_{i,j+1} - z_{i,j}| \leq \frac{1}{2R}(|z_{i,0} - a|^2 - |z_{i,j} - a|^2) + \alpha T_j \left(1 + \frac{1}{R}|z_{i,0} - a|\right) +$$

$$+\frac{\alpha^2}{2R}T_j^2+\frac{1}{2R}(|z_{i,j}-a|^2-|z_{i,j+1}-a|^2)+\alpha h\left(1+\frac{1}{R}|z_{i,0}-a|\right)+\frac{\alpha^2}{2R}(h^2+2hT_j)$$
$$=\frac{1}{2R}(|z_{i,0}-a|^2-|z_{i,j+1}-a|^2)+\alpha T_{j+1}\left(1+\frac{1}{R}|z_{i,0}-a|\right)+\frac{\alpha^2}{2R}T_{j+1}^2$$

where we have used the fact that $T_{j+1}=T_j+h$ and $T_{j+1}^2=T_j^2+h^2+2hT_j$. The same approach can be used to prove (106). The first two inequalities in (101) are then deduced. The third inequality follows from (94) from which one deduces that

$$||\xi_{k+1}-\xi_k||\leq(\gamma+\delta)\alpha h \tag{109}$$

The constant K in (102) can be obtained by summing the three variations in (101) and using the values for w_i and w_r. ■

Consider the step function $\mu^N(\cdot):[0,T]\to I\!R^r$ such that $\mu^N(t)=\mu_{k+1}$ for all $t\in[t_k,t_{k+1})$.

Proposition 13 *For any given $h^*>0$, there exists a constant $K'\equiv K'(h^*)>0$ such that:*

$$\mathit{var}(\mu^N,[0,T])\leq K',\forall h\in]0,h^*[. \tag{110}$$

Proof: Let $h^*>0$ be given. From (95) we deduce that

$$\mu_{k+1}-\mu_k=G^{-1}(I_r\ \ 0_{r\times(n-r)})(I_n-hWAW^{-1})(z_{k+1}-z_k)+G^{-1}(I_r\ \ 0_{r\times(n-r)})(z_k-z_{k-1}) \tag{111}$$

It follows that

$$\sum_{k=1}^{N-1}||\mu_{k+1}-\mu_k||\leq\|G^{-1}(I_r\ 0_{r\times(n-r)})(I_n-hWAW^{-1})\|\sum_{k=1}^{N-1}||z_{k+1}-z_k||+$$
$$+\|G^{-1}(I_r\ 0_{r\times(n-r)})\|\sum_{k=0}^{N-2}||z_{k+1}-z_k||\leq$$
$$\leq\|G^{-1}(I_r\ 0_{r\times(n-r)})(I_n-hWAW^{-1})\|\sum_{k=0}^{N-1}||z_{k+1}-z_k||+$$
$$+\|G^{-1}(I_r\ 0_{r\times(n-r)})\|\sum_{k=0}^{N-1}||z_{k+1}-z_k|| \tag{112}$$
$$\leq\|G^{-1}(I_r\ 0_{r\times(n-r)})\|(2+h^*\|WAW^{-1}\|)\sum_{k=0}^{N-1}||z_{k+1}-z_k|| \tag{113}$$

It follows from (102) that K' exists that depends on K and the system parameters, such that (110) is satisfied. ■

Remark 14 *Let us study the behaviour of $z_{1,k}$. One peculiarity of the inclusion* (92) *is that $T_\Phi^0 = I\!R^+$. Therefore $z_{1,k}$ =prox$[I\!R^+; z_{1,k-1} + hz_{2,k}] \geq 0$ for all $k \geq 1$. In other words $z_1^N(\cdot)$ may be negative only on $t \in [0, t_1)$. On the contrary, the other variables $z_i^N(\cdot)$ may become negative at any time $t \geq 0$. Consider for instance $T_\Phi^{i-1}(z_{1,k}, \ldots, z_{i-1,k}) = I\!R$ for all $2 \leq i \leq r$, so that $z_{2,k} = z_{2,k-1} + h\sum_{i=3}^r z_{i,k} + hCA^rW^{-1}z_k$. Nothing hampers that ξ_k takes a value such that $z_{2,k} < 0$ even if $z_{i,k-1} > 0$ for all $2 \leq i \leq r$. It follows from* (92), (93) *and convex analysis that*

$$0 \leq z_{1,k+1} \perp \mu_{1,k+1} \geq 0 \quad \text{for all } \ k \geq 0 \tag{114}$$

and

$$0 \leq z_{1,k+1} \perp \mu_{r,k+2} \geq 0 \quad \text{for all } \ k \geq 0 \tag{115}$$

which is the discrete-time complementarity condition corresponding to (54) and (58).

5.3.4 Convergence of the discretized solution

We now denote $\{z^N(\cdot)\}$ the sequence of functions constructed from the functions $z^N(\cdot)$, and similarly for $\mu^N(\cdot)$.

Proposition 14 *There exists a subsequence $\{z^{N_k}\}$ of $\{z^N\}$ which converges point-wisely to some function $z(\cdot) : [0,T] \to I\!R^n$, with variation $\leq K$, and a subsequence $\{\mu^{N_k}\}$ of $\{\mu^N\}$ which converges point-wisely to some function $\mu(\cdot) : [0,T] \to I\!R^r$, with variation $\leq K'$.*

Proof: The function $z^N(\cdot)$ is uniformly bounded on $[0,T]$, and it has bounded variation on $[0,T]$, see Propositions 11 and 12. From [45, Theorem 0.2.1 (i), (6)] the result follows. The proof is the same for $\mu^N(\cdot)$, using Propositions 11 and 13. ∎

Remark 15 *The convergence of $\mu^N(\cdot)$ towards a LBV* function *reflects the fact that, as said in the introduction of section 5.2, the primary variables are $\mu_{i,k+1} \triangleq d\nu_i((t_k, t_{k+1}])$. Hence the Dirac measures do not appear in the limit $\mu(\cdot)$ which is by construction a (bounded) function. The limits in Proposition 14 are unique.*

The next proposition establishes the ($*$-weak) convergence of the "derivatives" of the step functions $z_i^N(\cdot)$, which are Stieltjes measures, towards the Stieltjes measures dz_i.

Proposition 15 *For every continuous function of bounded variation $\varphi : [0,T] \to I\!R$ and all $1 \leq i \leq r$ we have:*

$$\int_{(s,t]} \varphi \, dz_i^{N_k} \to \int_{(s,t]} \varphi \, dz_i \ \ (s < t) \ \ \text{as } N_k \to +\infty \tag{116}$$

Proof: This is a consequence of [45, Theorem 0.2.1 (iii)] and the fact that the $z_i^N(\cdot)$ are right continuous functions. ∎

The differential measures $dz_i^{N_k}$ are of the form $dz_i^{N_k} = \{\dot{z}_i^{N_k}\}(t)dt + \sum_{k=1}^{N-1}(z_{i,k+1} - z_{i,k})\delta_{t_{N,k}}$(the singular part being zero since $z_i^{N_k}(\cdot)$ is a step function). Since the z_i's are LBV, dz_i admits a similar decomposition (see section 2), with atoms at times τ_k.

The proof that the limit functions are solutions of the time-continuous sweeping process (44)-(45)-(46) is left as a future work. Examples 5 and 6 in Section 5.5.1 demonstrates however the fact that the solutions of (92), (93) and (94) do converge to the higher oder sweeping process solutions in simple cases.

5.4 Overview of the implementation

In this section, we provide a short overview of the implementation of the numerical algorithm of the Extended Moreau Sweeping process. Our goal is to present the algorithm in a pedagogical way through a straightforward implementation. Obviously, for efficiency reasons, the code may slightly differ from what is below.

Matrix formulation of the ZD form in view of numerical integration. Given $W \in \mathbb{R}^{n\times n}$, the linear transformation of the state space, we introduce the following matrix notations :

$$[I - h\bar{A}]z_{k+1} = z_k + \bar{G}\mu_{k+1} \tag{117}$$

where the matrices $\bar{A} \in \mathbb{R}^{n\times n}$ and $\bar{G} \in \mathbb{R}^{n\times r}$ are defined as follows :

$$\bar{A} = \left[\begin{array}{c|c|c} \begin{matrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \end{matrix} & & 0 \\ \hline \multicolumn{3}{c}{CA^rW^{-1}} \\ \hline B_\xi & 0 & A_\xi \end{array}\right] \quad \bar{G} = \left[\begin{array}{c} \begin{matrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots\dots\dots & & 0 & CA^{r-1}B \end{matrix} \\ \hline 0_{(n-r)\times r} \end{array}\right] \tag{118}$$

Notice that $\bar{A} = WAW^{-1}$ is defined in (19) and that $\bar{G}$ is defined by (66).

Expression of the inclusions (91) **in terms of nested complementarity problems.** Let us consider the following inclusion :

$$\mu_{1,k+1} \in -\partial\psi_\Phi(z_{1,k+1}).$$

Let us set $\Phi = I\!R^+$. This inclusion may be stated equivalently as a complementarity problem[2] :

$$0 \leq z_{1,k+1} \perp \mu_{1,k+1} \geq 0$$

If $r > 1$ then we must handle the second inclusion :

$$\mu_{2,k+1} \in -\partial \psi_{T^1_\Phi(z_{1,k})}(z_{2,k+1})$$

which can be reformulated in terms of a complementarity problem :

$$\text{If } z_{1,k} \leq 0, \text{ then } 0 \leq z_{2,k+1} \perp \mu_{2,k+1} \geq 0$$

In this way, for $r > 2$, we get the following complementarity problem :

$$\text{If } z_{1,k} \leq 0 \text{ and } z_{2,k} \leq 0, \text{ then } 0 \leq z_{3,k+1} \perp \mu_{3,k+1} \geq 0$$

In the general case, we search the integer $1 \leq r^\star \leq r$ satisfying the following condition :

$$r^\star = \begin{cases} 1, \text{ if } z_{1,k} > 0 \\ 1 + \max\{j \leq r-1 : \quad z_{i,k} \leq 0, \forall i \leq j\} \end{cases}$$

Then we obtain the following set of nested complementarity problems :

$$0 \leq z_{i,k+1} \perp \mu_{i,k+1} \geq 0, \quad 1 \leq i \leq r^\star \tag{119}$$

We define the vectors collecting the state and the multiplier for the "active" constraints by :

$$\begin{aligned} z^\star_{k+1} &= [z_{1,k+1}, \ldots, z_{r^\star,k+1}]^T \\ \mu^\star_{k+1} &= [\mu_{1,k+1}, \ldots, \mu_{r^\star,k+1}]^T. \end{aligned} \tag{120}$$

and we introduce the matrix $R \in I\!R^{r^\star \times r}$ describing the relation between $z^\star_{k+1}$ and z_{k+1} :

$$z^\star_{k+1} = R\bar{z}_{k+1}. \tag{121}$$

Assuming that $\mu_{i,k+1} = 0, i > r^\star$, we get the relation between $\mu_{i,k+1}$ and $\mu^\star_{k+1}$:

$$\mu_{k+1} = R^T \mu^\star_{k+1}. \tag{122}$$

[2]In a more general setting, a cone complementarity problem has to be written $\Phi \ni z_{1,k+1} \perp -\mu_{1,k+1} \in \Phi^\star$, with $\Phi^\star$ the dual cone of Φ.

Formulation of the one-step LCP problem To be more explicit in the computation of the state vector, we introduce the matrix $P \in I\!R^{r \times n}$ such that

$$\bar{z}_{k+1} = P z_{k+1}. \tag{123}$$

Assuming that $r^\star$ is computed at each step and that z_k is known, the following set of discretized equations has to be solved to advance from step k to step $k+1$:

$$\begin{cases} [I - h\bar{A}] z_{k+1} = z_k + \bar{G}\mu_{k+1} \\ \bar{z}_{k+1} = P z_{k+1} \\ z^\star_{k+1} = R \bar{z}_{k+1} \\ \mu_{k+1} = R^T \mu^\star_{k+1} \\ 0 \leq \mu^\star_{k+1} \perp z^\star_{k+1} \geq 0. \end{cases} \tag{124}$$

This yields the following closed form for the one-step LCP problem :

$$\begin{cases} z^\star_{k+1} = RP[I - h\bar{A}]^{-1} z_k + RP[I - h\bar{A}]^{-1} \bar{G} R^T \mu^\star_{k+1} \\ 0 \leq \mu^\star_{k+1} \perp z^\star_{k+1} \geq 0. \end{cases} \tag{125}$$

Pseudo-algorithm for the EMTS A pseudo-algorithm for the EMTS is given in order to clearly outline the major features of the numerical resolution of the EMTS (see algorithm 1).

5.5 Numerical applications

5.5.1 Comparison with a Backward Euler Scheme

Several attempts have been already made to solve numerically dynamical systems with arbitrary relative degree. In [30], an algorithm for constructing solutions rather than a pure numerical scheme, is given based on an event-driven strategy. This type of strategy cannot encompass general evolutions with accumulation of events, and is not well suited for convergence proofs. In [17], the direct use of a backward Euler scheme with an implicit evaluation of the complementarity condition yields a time-stepping scheme, which is very similar to the "catching up algorithm" of Moreau. This scheme works well with systems of relative degree less or equal to 1, but exhibits characteristic examples of inconsistency for higher relative degree. We have collected in this section some remarks on the behaviour of the numerical scheme presented above which lead us to believe that the EMTS scheme solves this problem.

For the first and the second order sweeping process, the time integration method is often confused with a standard backward Euler scheme. To highlight the difference with the numerical time integration of the Moreau's sweeping process, we consider several examples of inconsistencies introduced in [17]. A naive way of integrating a LCS is to apply directly a backward Euler scheme:

$$\begin{cases} \dfrac{x_{k+1}-x_k}{h} = Ax_{k+1} + B\lambda_{k+1} \\[2ex] w_{k+1} = Cx_{k+1} + D\lambda_{k+1} \\[2ex] 0 \leq \lambda_{k+1} \perp w_{k+1} \geq 0 \end{cases} \tag{126}$$

which can be reduced to a LCP by a straightforward substitution:

$$0 \leq \lambda_{k+1} \perp C(I-hA)^{-1}x_k + hC(I-hA)^{-1}B\lambda_{k+1} \geq 0 \tag{127}$$

In the sequel, such a LCP will be denoted as $(y_{k+1}, \lambda_{k+1}) = LCP(M, b_{k+1})$ where

$$M = hC(I-hA)^{-1}B \tag{128}$$

$$b_{k+1} = C(I-hA)^{-1}x_k \tag{129}$$

In [17], some consistency and convergence results are proved. Shortly, under the assumption that D is nonnegative definite or that the triplet (A, B, C) is a minimal representation and (A, B, C, D) is positive real (see appendix C), they exhibit that some subsequences of $\{y_k\}, \{\lambda_k\}, \{x_k\}$ converge weakly to a solution y, λ, x of the LCS. Such assumptions imply that the relative degree r is less or equal to 1. In the case of the relative degree 0, the LCS is equivalent to a standard system of ordinary differential equations with a Lipschitz-continuous vector field (see [24], Remark 10). The result of convergence is then the standard result of convergence for the Euler backward scheme. In the case of a relative degree equal to 1, these results corroborate the results of [9].

As we said earlier, several examples for which the backward Euler scheme does not work at all are also detailed in [17]. These systems are of higher relative degree. We will consider below two similar examples and comment on the difference between the Backward Euler scheme and our approach.

Example 5 *Let us consider a LCS with the following matrix definition:*

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}; B = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}; C = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}; D = 0 \tag{130}$$

The relative degree r of this LCS is equal to 2 ($D = 0, CB = 0, CAB \neq 0$). If we apply the time discretization given by (126)*, we can remark that:*

$$\lim_{h \longrightarrow 0} \frac{1}{h} M = \lim_{h \longrightarrow 0} C(I-hA)^{-1}B = 0 \tag{131}$$

It is clear that if h is taken very small, which may be needed in many practical cases or for the convergence analysis of the scheme, then the LCP matrix in (127) has little chance to be well conditioned due to the fact that $CB = 0$.

If we consider the initial data $x_0 = (0, -1, 0)^T$, we obtain by a straightforward application of the scheme (126) *the following solution:*

$$x_k = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} ; \forall k \geq 1 \tag{132}$$

$$\lambda_1 = \frac{1}{h}; \quad \lambda_k = 0, \forall k \geq 2 \tag{133}$$

We can remark that the multiplier λ_1 which is the solution of the LCP at the first step, tends towards $+\infty$ when h vanishes. In this example, the state $x(t)$ seems to be well approximated but both the LCP matrix and the multiplier tend to inconsistent value when h vanishes. This inconsistency is just the result of an attempt to approximate the point value of a distribution, which is nonsense.

If we consider now the initial data $x_0 = (-1, -1, 0)^T$, we obtain the following numerical solution from (126) :

$$x_k = \begin{pmatrix} k \\ \frac{1}{h} \\ 0 \end{pmatrix} ; \forall k \geq 1 \tag{134}$$

$$\lambda_1 = \frac{1}{h^2}; \quad \lambda_k = 0, \forall k \geq 2 \tag{135}$$

With such an initial data, the exact solution should be $x_k = 0, \forall k \geq 1$. We can see that there is an inconsistency in the result because the first component of the approximate state does not depend on the time-step. We can not expect that this approximation converges to the exact solution.

If we apply the EMTS scheme, we obtain the following solution:

$$x_k = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} ; \forall k \geq 1 \tag{136}$$

$$\mu_{1,1} = 1; \quad \mu_{2,1} = 1, \tag{137}$$

$$\mu_{1,k} = 0, \mu_{2,k} = 0, \ \forall k \geq 2 \tag{138}$$

which converges to the time-continuous solution of the higher order Moreau's sweeping process, i.e. $x(0) = x_0, x(t) = (0, 0, 0)^T, \forall t > 0$.

Example 6 *Let us consider this second very simple example:*

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} ; B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} ; C = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} ; D = 0 \tag{139}$$

In this case, the relative degree, r is equal to 3. The direct discretization of the system leads to the same problem as in the previous example even in the case where the initial data satisfies the constraints. Let us consider the case of $x_0 = (0, -1, 0)^T$. From (126), we obtain the following numerical solution:

$$x_k = \begin{pmatrix} \dfrac{k(k+1)}{2h} \\ k \\ \dfrac{1}{h} \end{pmatrix}; \forall k \geq 1 \tag{140}$$

$$\lambda_1 = \frac{1}{h^2}; \quad \lambda_k = 0, \forall k \geq 2 \tag{141}$$

This solution can not converge to an analytical solution. The solution given by the EMTS scheme is

$$x_k = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \forall k \geq 1 \tag{142}$$

$$\mu_{1,1} = 1; \quad \mu_{2,1} = 1, \quad \mu_{3,1} = 0, \tag{143}$$

$$\mu_{i,k} = 0, \mu_{2,k} = 0, \ \forall k \geq 2, i = 1 \ldots 3 \tag{144}$$

which is the time-continuous solution of the higher order Moreau's sweeping process, i.e. $x(0) = x_0, x(t) = (0, 0, 0)^T, \forall t > 0$.

Interest of the ZD canonical form from the numerical point of view. Let us consider now the ZD form with $r = n$ for simplicity sake. A direct discretization with the Euler backward scheme leads to:

$$\begin{cases} \dfrac{z_{k+1} - z_k}{h} = \bar{A} z_{k+1} + \hat{B} \lambda_{k+1} \\ \\ 0 \leq z_{1,k+1} \perp \lambda_{k+1} \geq 0 \end{cases} \tag{145}$$

with $\hat{B} = (0, \ldots, 0, CA^{r-1}B)^T$. Using (145), we get the following LCP $(z_{1,k+1}, \Lambda_{k+1}) = LCP(\frac{1}{h^r} M, b_{k+1})$ with the LCP matrix given by

$$M = CA^{r-1}[hW(I - h\bar{A})^{-1}\hat{B} + B]. \tag{146}$$

It is clear from (146) that if h vanishes, then the LCP matrix is close to the matrix $CA^{r-1}B$ which is the LCP matrix of the time-continuous ZD form. So if $CA^{r-1}B$ assures that the LCP(λ) has a unique solution, this should be the case for the discretized LCP as well, for small enough step h.

The interest of working with the ZD dynamics lies in the fact that this allows one to keep the properties of the LCP from the continuous time t formulation to the discretized formulation.

This seems to be some kind of minimal requirement for the discrete algorithm, since in any case λ_{k+1} has to be calculated[3].

5.5.2 Influence of the zero-dynamics on the solution

In this example, we illustrate how the zero dynamics may influence the behaviour of the following system

$$\left\{\begin{array}{l} \dot{z}_1(t) = z_2(t) \\ \dot{z}_2(t) = z_3(t) \\ \dot{z}_3(t) = -z_1(t) - z_2(t) - z_3(t) - d_\xi^T \xi(t) + \lambda(t) \\ \dot{\xi}_1(t) = \alpha \xi_2(t) \\ \dot{\xi}_2(t) = -\omega \xi_1(t) + z_1(t) \\ w(t) = z_1(t) \geq 0 \end{array}\right. \tag{147}$$

with the initial condition $z(0) = (1,0,0,0,0)^T$. The system (147) of relative degree $r = 3$, is embedded in the MDI formalism (44)-(46). All the simulations are performed with Scilab$^{\copyright}$. The time interval is $[0,10]$ and the time step is equal to $h = 10^{-1}$.

In the first experiment, we choose $d_\xi = (0,0), \alpha = 1$ and $\omega = 1$. The zero dynamics, which is a pure harmonic oscillator, does not play any role in the $\bar{z}$-dynamics. The results are depicted on Figure 2. The system seems to be stable and the state vanishes.

Choosing $d_\xi = (0,1), \alpha = 1$ and $\omega = 1$, the same simulation is made now with a coupled zero dynamics. The influence of the zero dynamics is shown on Figure 3.

If we choose $d_\xi = (0,1)$ but $\alpha = 0$ and $\omega = 0$, $\bar{z}(t)$ is identically equal to zero after the first jump. The simulation is shown on Figure 4. Finally, if we choose $d_\xi = (0,-1)$ but $\alpha = 1$ and $\omega = 1$, there are non-zero intervals on which the constraints remain active, see Figure 5.

It is therefore clear that the zero-dynamics matrix A_ξ and the connection vector d_ξ, have a strong influence on the system's dynamics. In particular, one sees that despite the state reinitialization mapping sets the post-impact states $\bar{z}$ to zero (plastic impacts), the zero dynamics may force the system to detach from the constraint and undergo subsequent jumps after a boundary trajectory has occurred.

5.5.3 Empirical order of the scheme.

This section is devoted to provide an empirical estimation of the order of the scheme. Let us consider the previous example with $d_\xi = (0,-1), \alpha = 1, \omega = 1$ and with the initial condition $z(0) = (1,0,0,0,0)^T$.

In order to evaluate the order of accuracy of the scheme on this simple example, we need to use a norm which is consistent with the set of RCLBV functions and to introduce a notion of convergence providing a reasonable substitute to the uniform convergence of the

[3] Some schemes and some dynamical formulation, do not use an explicit calculation of the multiplier. But they necessarily use underlying arguments equivalent to having a well-posed LCP.

continuous functions. To overcome this difficulty, the convergence in the sense of filled-in graph has been introduced by Moreau [52]. Shortly, for a RCLBV function $f : [0,T] \mapsto \mathbb{R}^n$, we define the filled-in graph, $gr^\star f$ by adding some line segments to the graph of f in such a way that all the gaps are filled:

$$gr^\star f = \{(t,x) \in [0,T] \times \mathbb{R}^n, 0 \leq t \leq T \text{ and } x \in [f(t^-), f(t^+)]\} \tag{148}$$

Such graphs are closed bounded subsets of $[0,T] \times \mathbb{R}^n$, hence, we can use the Hausdorff distance between two such sets with a suitable metric:

$$d((t,x),(s,y)) = \max\{|t-s|, \|x-y\|_{\mathbb{R}^n}\} \tag{149}$$

Defining the excess of separation between two graphs by

$$e(gr^\star f, gr^\star g) = \sup_{(t,x)\in gr^\star f} \inf_{(s,y)\in gr^\star g} d((t,x),(s,y)), \tag{150}$$

the Hausdorff distance between two filled-in graphs $h^\star$ is defined by

$$h^\star(gr^\star f, gr^\star g) = \max\{e(gr^\star f, gr^\star g), e(gr^\star g, gr^\star f)\} \tag{151}$$

To compute a reference solution, the number of time-steps is chosen as $N = 10^6$, i.e., for a time step $h = 10^{-5}$. The error with the norm of the uniform convergence $\|\|_\infty$ is displayed in log scale on the Figure 6. We can see that there is no way to measure the rate of convergence with the $\|\|_\infty$ norm. The results of the distance in the sense of filled-in graph is displayed in log scale on the Figure 7. On this example, the order of accuracy of the EMTS scheme is close to 1, as expected.

6 Applications: Electrical circuits with ideal diodes

The well-posedness of dissipative circuits with ideal diodes ($r = 0$ or $r = 1$) has been investigated in [18, 12] and their time-discretization with time-stepping Euler implicit schemes is studied in [17]. In this section we briefly illustrate how the application of feedback signals in simple electrical circuits may lead to higher relative degree complementarity systems. In other words we show how the material of the foregoing sections may help in understanding the closed-loop dynamics of some complementarity systems.

A single input/single output case

Let us consider the following dynamics that corresponds to the circuit depicted in figure 8:

$$\left\{\begin{array}{l} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = -\frac{R}{L}x_2(t) + \frac{u(t)}{L} - \frac{1}{LC}x_1(t) - \frac{1}{L}\lambda(t) \\ \\ 0 \leq \lambda(t) \perp w(t) = -x_2(t) \geq 0 \end{array}\right. \tag{152}$$

where $x_2(\cdot)$ is the current across the circuit and $-\lambda$ is the voltage of the diode, $u(\cdot)$ is a voltage control. If $u(\cdot) = 0$ then one sees that the transfer function of the operator $\lambda(\cdot) \mapsto w(\cdot)$ is positive real, equal to $\frac{LCs}{LCs^2+RCs+1}$. Let us apply the following feedback

$$\left\{\begin{array}{l} u(t) = \lambda(t) + Lx_3(t) \\ \\ \dot{x}_3(t) = x_4(t) \\ \dot{x}_4(t) = \lambda(t). \end{array}\right. \tag{153}$$

Inserting (153) into (152) one obtains

$$\left\{\begin{array}{l} \dot{x}_1(t) = x_2(t) \\ \dot{x}_2(t) = -\frac{R}{L}x_2(t) - \frac{1}{LC}x_1(t) + x_3(t) \\ \dot{x}_3(t) = x_4(t) \\ \dot{x}_4(t) = \lambda(t) \\ \\ 0 \le w(t) = -x_2(t) \end{array}\right. \iff \left\{\begin{array}{l} \dot{z}_1(t) = z_2(t) \\ \dot{z}_2(t) = z_3(t) \\ \dot{z}_3(t) = \frac{R}{L}z_3(t) + \frac{1}{LC}z_2(t) + \lambda(t) \\ \dot{\xi}(t) = z_1(t) \\ \\ 0 \le w(t) = z_1(t) \end{array}\right. \tag{154}$$

We embed the ZD dynamics in (154) in the higher order sweeping process formalism. The relative degree is $r = 1$ in (152), but applying a dynamic feedback changes it since $r = 3$ in (154), and $CA^2B = 1$. It is immediate that $r - 1$ is equal to the number of integrations in the control input in (153). It is noteworthy that one can apply Proposition 7. In other words, the feedback law in (153) augments the relative degree, but does not destroy the dissipativity of the closed loop system, since the transfer function of the operator $\lambda \mapsto z_3$ in (154) is positive real. This puts electrical circuits in a perspective that perfectly fits with the higher order sweeping process.

Here $CA^2B = 1 > 0$. We can thus apply Corollary 2 to assert that for each $(\bar{z}_0, \xi_0) \in \mathbb{R}^n$, the system in (154) has a unique regular solution.

Remark 16 **i)** *We do not wish to discuss here the physical applicability of the feedback law in (153) (in practice measuring the voltage of the diode may introduce further dynamics), nor of an ideal diode model. At this stage it is however fundamental to keep in mind that studying such models allows the designer to point out and understand some phenomena which may be only limits (in the mathematical sense) of the real phenomena, but which would have been hidden by any other sort of modeling approach like penalisation. It is also worth recalling that our model has some intrinsic flexibility, see e.g. remark 12, and may therefore be adapted to better fit with the physical observations.*

ii) *Optimal control under state constraints also yields higher order complementarity systems [27, 4, 74]. Consequently it may benefit from the work in this paper.*

7 Conclusion

In this paper we present an extension of Moreau's sweeping process, a widely studied differential inclusion in the field of unilateral Mechanics. This provides a new formalism for higher relative degree complementarity systems. It allows us to 1) obtain a clear understanding of the dynamical mechanism which permits the integration of such systems where higher degree distributions naturally appear, 2) to derive a numerical time-stepping scheme for Initial Value Problems. This paper focuses on the formalism and time-discretization aspects. The dynamical framework that is presented possesses several interesting features:

- The formalism easily extends to nonlinear, time varying vector fields, time varying and state dependent sets $\Phi(t, z)$, and the state re-initialization law can be modified (consequently it encompasses Lagrangian systems),
- the sweeping process differential inclusion is a suitable formalism to design time-stepping numerical schemes, which paves the way towards convergence analysis thanks to its compact formulation as a discretized differential inclusion,
- further extensions and their analysis may benefit a lot from the numerous studies on the first and second order sweeping processes,
- preliminary well-posedness results show that the framework is sound,
- important potential application fields like electrical circuits, optimal control with state constraints (which is itself quite related to the dual problem of the so-called Continuous-Time Linear Programming problem [60, 3] which inherently involves distributional solutions), may benefit from the approach.

This paper brings some elements of answer to a questioning in [59], about the interpretation of the existence theory for differential variational inequalities with index ≥ 3 (i.e. $r \geq 2$ in the language of this paper).

Acknowledgements

This work is partially supported by the European project SICONOS IST2001-37172. We thank Professor Anstreicher for pointing out the references [60] and [3].

References

[1] V. Acary, B. Brogliato, A. Daniilidis, and C. Lemaréchal. On the equivalence between complementarity systems, projected systems and unilateral differential inclusions. *Systems and Control Letters, to appear*, 2005. Preprint INRIA Research Report RR-5107, available at http://www.inria.fr/rrrt/index.en.html.

[2] S. Adly and D. Goeleven. A stability theory for second order nonsmooth dynamical systems with application to friction problem. *Journal de Mathématiques Pures et Appliquées*, 83:17–51, 2004.

[3] K.M. Anstreicher. Generation of feasible descent directions in continuous time linear programming. Technical Report SOL 83-18, Systems Optimization Laboratory, Dept. of Operations Research, Stanford University, 1983.

[4] A.V. Arutyunov and S.M. Aseev. State constraints in optimal control. the degeneracy phenomenon. *Systems and Control Letters*, 26:267–273, 1995.

[5] J.P. Aubin and A. Cellina. *Differential Inclusions.* Springer Verlag, Berlin, 1994.

[6] P. Ballard. The dynamics of discrete mechanical systems with perfect unilateral constraints. *Archives for Rational Mechanics and Analysis*, 154:199–274, 2000.

[7] H. Benabdellah, C. Castaing, A. Salvadori, and Syam. A. Nonconvex sweeping process. *Journal of Applied Analysis*, 2(2):217–240, 1996.

[8] A. Berman and Plemmons. *Nonnegative Matrices in The Mathematical Sciences.* Computer Science and Applied Mathematics. academic Press, New York, 1979.

[9] H. Brezis. *Opérateurs maximaux monotones et semi-groupe de contraction dans les espaces de Hilbert.* North Holland, Amsterdam, 1973.

[10] B. Brogliato. *Nonsmooth Mechanics.* Springer Verlag London, second edition, 1999.

[11] B. Brogliato. Some perspectives on the analysis and control of complementarity systems. *IEEE Transactions on Automatic Control*, 48(6):918–935, 2003.

[12] B. Brogliato. The absolute stability problem and the Lagrange-Dirichlet theorem with monotone multivalued mappings. *Systems and Control Letters*, 51(5):343–353, 2004.

[13] B. Brogliato. Some results on the controllability of planar variational inequalities. *Systems and Control Letters*, 54:65–71, 2005.

[14] B. Brogliato, S. Niculescu, and P. Orhant. On the control of finite-dimensional mechanical systems with unilateral constraints. *IEEE Transactions on Automatic Control*, 42(2):200–215, 1997.

[15] B. Brogliato, A.A. ten Dam, L. Paoli, F. Genot, and M. Abadie. Numerical simulation of finite dimensional multibody nonsmooth mechanical systems. *ASME Applied Mechanics Reviews*, 55(2):107–150, 2002.

[16] H. Brézis. Problèmes unilatéraux. *Journal de Mathématiques pures et appliquées*, 51:1–168, 1972.

[17] K. Camlibel, W.P.M.H. Heemels, and J.M. Schumacher. Consistency of a time-stepping method for a class of piecewise-linear networks. *IEEE Trans. Circuits and systems I*, 49:349–357, 2002.

[18] K. Camlibel, W.P.M.H. Heemels, and J.M. Schumacher. On linear passive complementarity systems. *European Journal of Control*, 8(3):220–237, 2002.

[19] C. Castaing, T.X. Duc Ha, and M. Valadier. Evolution equations governed by the sweeping process. *Set-Valued Analysis*, 1:109–139, 1993.

[20] C. Castaing and M.D.P. Monteiro-Marques. Evolution problems associated with nonconvex closed moving sets with bounded variation. *Portugaliae Mathematica*, 1:73–87, 1996.

[21] D. Cobb. On the solution of linear differential equations with singular coefficients. *Journal of Differential Equations*, 1982.

[22] L.C. Evans and R.F. Gariepy. *Measure Theory and Fine Properties of Sobolev Functions*. Studies in Advanced Mathematics. Boca Raton, 1992.

[23] F. Facchinei and J.-S. Pang. *Finite-dimensional variational inequalities and complementarity problems*, volume I & II of *Springer Series in Operations Research*. Springer Verlag NY. Inc., 2003.

[24] D. Goeleven and B. Brogliato. Stability and instability matrices for linear evolution variational inequalities. *IEEE Transactions on Automatic Control*, 49(4):521–534, 2004.

[25] D. Goeleven, D. Motreanu, Y. Dumont, and M. Rochdi. *Variational and Hemivariational Inequalities: Theory, Methods and Applications; Volume I: Unilateral Analysis and Unilateral Mechanics*. Nonconvex Optimization and its Applications. Kluwer Academic Publishers, 2003.

[26] D. Goeleven, D. Motreanu, and V.V. Motreanu. On the stability of stationary solutions of first order evolution variational inequalities. *Advances in Nonlinear Variational Inequalities*, 6:1–30, 2003.

[27] R.F. Hartl, S.P. Sethi, and R.G. Vickson. A survey of the maximum principles for optimal control problems with staet constraints. *S.I.A.M. Reviews*, 37:181–218, 1995.

[28] W.P.M.H. Heemels. *Linear Complementarity Systems. A Study in Hybrid Dynamics*. PhD thesis, Technical University of Eindhoven, 1999. ISBN 90-386-1690-2.

[29] W.P.M.H. Heemels and B. Brogliato. The complementarity class of hybrid dynamical systems. *European Journal of Control*, 9:311–349, 2003.

[30] W.P.M.H. Heemels, J.M. Schumacher, and S. Weiland. Linear complementarity problems. *S.I.A.M. Journal of applied mathematics*, 60(4):1234–1269, 2000.

[31] J.I. Imura. Well-posedness analysis of switch-driven piecewise affine systems. *IEEE Transactions on Automatic Control*, 2003.

[32] M. Jean. The non smooth contact dynamics method. *Comput. Methods. Appl. Mech. Eng.*, 177:235–257, 1999. Special issue on computational modeling of contact and friction, J.A.C. Martins and A. Klarbring, editors.

[33] A. Juloski, W.P.M.H. Heemels, and B. Brogliato. Observer design for lur'e systems with multivalued mappings. In *IFAC world Congress Prag*, 4–8 July 2005.

[34] P. Krejci and A. Vladimorov. Polyhedral sweeping processes with oblique reflection in the space of regulated functions. *Set-Valued Analysis*, 11:91–110, 2003.

[35] M. Kunze and M.D.P. Monteiro Marquès. Yosida-Moreau regularisation of sweeping processes with unbounded variation. *Journal of Differential Equations*, 130:292–306, 1996.

[36] M. Kunze and M.D.P. Monteiro Marquès. Existence of solutions for degenerate sweeping processes. *Journal of Convex Analysis*, 4:165–176, 1997.

[37] M. Kunze and M.D.P. Monteiro Marquès. On parabolic quasi-variational inequalities and state-dependent sweeping processes. *Topol. Methods Nonlinear Anal.*, 12:179–191, 1998.

[38] M. Kunze and M.D.P. Monteiro Marquès. An introduction to Moreau's sweeping process. In B. Brogliato, editor, *Impact in Mechanical systems: Analysis and Modelling*, volume 551 of *Lecture Notes in Physics*, pages 1–60. Springer, 2000.

[39] W.Q. Liu, W.Y. Lan, and K.L. Teo. On initial instantaneous jumps of singular systems. *IEEE Transactions on Automatic Control*, 1995.

[40] Y.J. Lootsma, A.J. van der Schaft, and K. Camlibel. Uniqueness of solutions of relay systems. *Automatica*, 35(3):467–478, 1999.

[41] R. Lozano, B. Brogliato, O. Egeland, and B. Maschke. *Dissipative Systems Analysis and Control. Theory and Applications*. Communications and Control Engineering. Springer Verlag, London, 2000.

[42] P. Lötstedt. Mechanical systems of rigid bodies subject to unilateral constraints. *SIAM Journal of Applied Mathematics*, 42(2):281–296, 1982.

[43] M. Mabrouk. A unified variational for the dynamics of perfect unilateral constraints. *European Journal of Mechanics - A/Solids*, 17, 1998.

[44] P. Machanda and A.H. Siddiqi. Rate dependent evolution quasivariational inequalities and state-dependent sweeping processes. *Advances Nonlinear Variational Inequalities*, 5(1):1–16, 2002.

[45] M. P. D. Monteiro Marques. *Differential Inclusions in NonSmooth Mechanical Problems : Shocks and Dry Friction.* Birkhauser, Verlag, 1993.

[46] J.J. Moreau. Les liaisons unilatérales et le principe de Gauss. *Comptes Rendus de l'Académie des Sciences*, 256:871–874, 1963.

[47] J.J Moreau. Quadratic programming in mechanics: dynamics of one sided constraints. *S.I.A.M. Journal on control*, 4(1):153–158, 1966.

[48] J.J. Moreau. Rafle par un convexe variable (première partie), exposé no 15. *Séminaire d'analyse convexe, University of Montpellier*, page 43 pages, 1971.

[49] J.J. Moreau. Rafle par un convexe variable (deuxième partie) exposé no 3. *Séminaire d'analyse convexe, University of Montpellier*, page 36 pages, 1972.

[50] J.J. Moreau. Problème d'évolution associé à un convexe mobile d'un espace hilbertien. *C.R. Acad. Sc. Paris, Série A-B,*, t.276:791–794, 1973.

[51] J.J. Moreau. Evolution problem associated with a moving convex set in a Hilbert space. *Journal of Differential Equations*, 26:347–374, 1977.

[52] J.J. Moreau. Approximation en graphe d'une évolution discontinue. *RAIRO Analyse numérique/ Numerical Analysis*, 12:75–84, 1978.

[53] J.J. Moreau. Liaisons unilatérales sans frottement et chocs inélastiques. *Comptes Rendus de l'Académie des Sciences*, 296 serie II:1473–1476, 1983.

[54] J.J. Moreau. Unilateral contact and dry friction in finite freedom dynamics. In J.J. Moreau and P.D. Panagiotopoulos, editors, *Nonsmooth mechanics and applications*, number 302 in CISM, Courses and lectures, pages 1–82. Springer Verlag, 1988.

[55] J.J. Moreau. Some numerical methods in multibody dynamics: Application to granular materials. *European Journal of Mechanics - A/Solids*, supp.(4):93–114, 1994.

[56] J.J. Moreau. Numerical aspects of the sweeping process. *Computer Methods in Applied Mechanics and Engineering*, 177:329–349, 1999. Special issue on computational modeling of contact and friction, J.A.C. Martins and A. Klarbring, editors.

[57] J.J. Moreau. An introduction to unilateral dynamics. In M. Frémond and F. Maceri, editors, *Novel Approaches in Civil Engineering*, Series: Lecture Notes in Applied and Computational Mechanics. Springer Verlag, 2003.

[58] K.G. Murty. *Linear and Nonlinear Programming.* Heldermann, 1988. available at http://www-personal.engin.umich.edu/~murty/book/LCPbook/.

[59] J.S. Pang and Stewart. Differential variational inequalities. preprint, 2004.

[60] A. Perold. Fundamentals of a continuous time simplex method. Technical Report SOL 78-26, Systems Optimization Laboratory, Dept. of Operations Research, Stanford University, 1978.

[61] A. Y. Pogromsky, W.P.M.H. Heemels, and H. Nijmeijer. On solution concepts and well-posedness of linear relay systems. *Automatica*, 39(12):2139–2147, 2003.

[62] R.J.B. Rockafellar, R.T.and Wets. *Variational Analysis*, volume 317. Springer Verlag, New York, 1997.

[63] M. Schatzman. Sur une classe de problèmes hyperboliques non linéaires. *Comptes Rendus de l'Académie des Sciences Série A*, 1973.

[64] M. Schatzman. A class of nonlinear differential equations of second order in time. *Nonlinear Analysis, Theory, Methods & Applications*, 2(3):355–373, 1978.

[65] L. Schwartz. *Analyse III, Calcul Intégral.* Hermann, 1993.

[66] G.E. Shilov and B.L. Gurevich. *Integral Measure and Derivative. A Unified Approach.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1966. Hermann, Paris, 1993.

[67] A.H. Siddiqi and P. Machanda. Variants of Moreau's sweeping process. *Adv. Nonlinear Var. Inequal.*, 5(1):17–28, 2002.

[68] A.H. Siddiqi, P. Machanda, and Brokate. M. On some recent developments concerning Moreau's sweeping process. In A.H. Siddiqi and M. Kocvara, editors, *International Conference on Emerging Areas in Industrial and Applied Mathematics,*. GNDU Amritsar, India, Kluwer Academic Publishers, Boston-London-Dordrecht, 2001, 2001.

[69] D. Stewart. Convergence of a time-stepping scheme for rigid-body dynamics and resolution of Painlevé's problem. *Archives for Rational Mechanics and Analysis*, 145:215–260, 1998.

[70] D. Stewart. Rigid body dynamics with friction and impact. *SIAM Reviews*, 42(1):3–39, 2000.

[71] D. Stewart. Reformulations of measure differential inclusions and their closed graph property. *Journal of Differential Equations*, 175(1):108–129, 2001.

[72] A.A. ten Dam, E. Dwarshuis, and J.C. Willems. The contact problem for linear continuous-time dynamical systems: a geometric approach. *IEEE Transactions on Automatic Control*, 42(4):458–472, 1997.

[73] L. Thibault. Sweeping process with regular and nonregular sets. *Journal of Differential Equations*, 193(1):1–26, 2003.

[74] A. J. van der Schaft and J.M. Schumacher. *An Introduction to Hybrid Dynamical Systems.* Springer Verlag London, 2000.

[75] A.J. van Der Schaft and J.M. Schumacher. The complementary-slackness class of hybrid systems. *Mathematics of Control, Signals and Systems*, 9(3):266–301, 1996.

[76] A.J. van der Schaft and J.M. Schumacher. Complementarity modeling of hybrid systems. *IEEE Transactions on Automatic Control*, 43(4):483–490, 1998.

[77] M. Vidyasagar. *Nonlinear Systems Analysis.* Prentice Hall, second edition, 1993.

A Lemma 0.4.3 in [45]

Lemma 4 *In a Hilbert space H, we consider a closed convex set C which contains a close ball $\bar{B}(a,R)$, $R>0$. Let $x \in H$. Then*

$$||x - \textit{prox}\,[C;x]|| \leq \frac{1}{2R}(||x-a||^2 - ||\textit{prox}\,[C;x]-a||^2) \tag{155}$$

B Lemma 0.4.4 in [45]

Lemma 5 *Let $C_1,\ldots,C_n$ be closed convex subsets of H, all of them containing a fixed closed ball $\bar{B}(a,R)$, $R>0$. If $x_0 \in H$ and if $x_1,\ldots,x_n$ are defined inductively by x_i =prox $[C_i;x_{i-1}]$, then:*

$$\begin{array}{l} ||x_0-a|| \geq ||x_1-a|| \geq \ldots \geq ||x_n-a|| \\ \sum_{i=1}^{n} ||x_i - x_{i-1}|| \leq \frac{1}{2R}(||x_0-a||^2 - ||x_n-a||^2) \leq \frac{1}{2R}||x_0-a||^2 \end{array} \tag{156}$$

C Positive real systems

Let a system be given by the following state space representation

$$\left\{ \begin{array}{l} \dot{x}(t) = Ax(t) + Bu(t) \\ \\ w(t) = Cx(t) \end{array} \right. \tag{157}$$

with $x \in I\!R^n$, $u \in I\!R^m$, $w \in I\!R^m$. We assume that (A,B) is controllable and that (A,C) is observable.

Definition 4 *The transfer function matrix $H(s) = C(sI_n - A)^{-1}B$, $s \in \mathbb{C}$, is said positive real if*

- *$H(s)$ is analytic in* $\mathbf{Re}[s] > 0$,
- *the smallest eigenvalue of $H(j\omega) + H^*(j\omega)$ is nonnegative for all $\omega \geq 0$.*

where $H^*(s)$ is the conjugate transpose of $H(s)$.

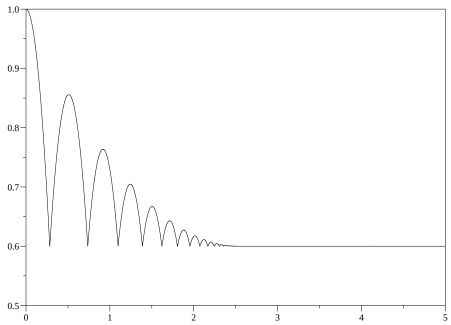
The following is called the Kalman-Yakubovic-Popov lemma [41].

Lemma 6 *The transfer function matrix* $H(s) = C(sI_n - A)^{-1}B$ *is positive real if and only if there exists matrices* $P = P^T > 0$, $L \in \mathbb{R}^{n\times m}$, *such that*

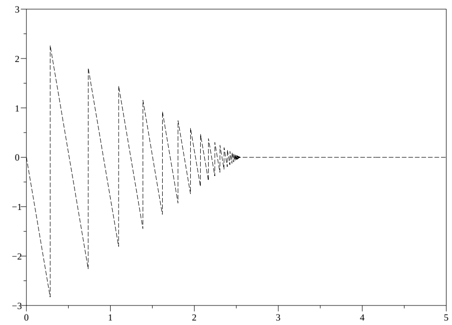
$$\begin{aligned} &PA + A^T P = -LL^T \\ &PB = C^T \end{aligned} \tag{158}$$

Algorithm 1 Sketch of the Extended Moreau's Time Stepping(EMTS) Scheme

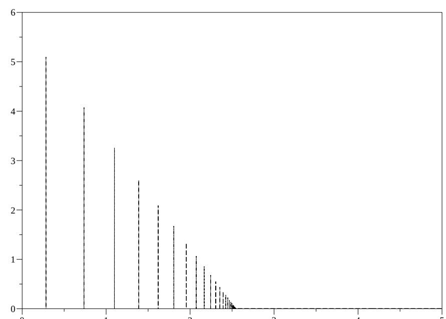
Require: Classical form of the system : A, B, C, x_0
Require: Zero-Dynamic form : r, W, A_ξ, B_ξ
Require: Numerical parameters h, T
Ensure: $(\{x_n\}, \{z_n\}, \{\mu_n\}) = Approx(A, B, C, D, x_0, h, T)$
  // Computation of the operator associated with the ZD form :
  $\bar{A}, \bar{G}, P$
  // Time discretization $N := [\frac{T}{h}]$
  // Computation of the time invariant numerical operators:
  $\bar{M} := (I - h\bar{A})^{-1}$
  $M_{lcptmp} := P\bar{M}\bar{G}$
  $b_{tmp} := P\bar{M}$
  $z_0 := Wx_0$
  // Loop in time.
  for $k = 0$ to N **do**
    //Compute the rank $r^\star$
    $r^\star = 1, i = 1;$
    while $z_k(i) \leq 0$ **do**
      $r^\star = r^\star + 1$
      $i = i + 1;$
    end while
    // Computation of R
    // Solve the one-step LCP problem
    $M_{lcp} := RM_{lcptmp}R^T$
    $b := b_{tmp}z_k$
    $(\mu_{k+1}, z^\star_{k+1}) := \mathsf{SolveLCP}(\mathsf{M_{lcp}}, \mathsf{b})$
    // State Actualization
    $z_{k+1} := \bar{M}\left[z_k + \bar{G}\mu_{k+1}\right]$
    $x_{k+1} := W^{-1}z_{k+1}$
  end for



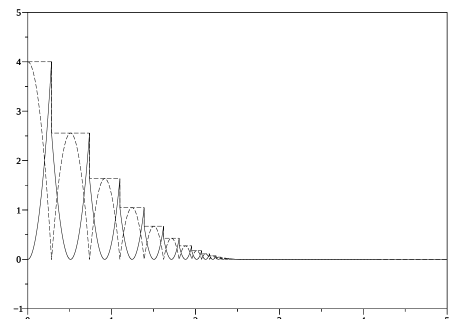
(a) Position of the ball vs. Time.



(b) Velocity of the ball vs. Time.



(c) Amplitude of the impulse vs. Time.



(d) Total, kinetic and potential vs. Time.

Figure 1: Bouncing Ball on a rigid plane. $e = 0.8, g = 10m.s^{-2}, m = 1kg, h = 5.10^{-3}s$

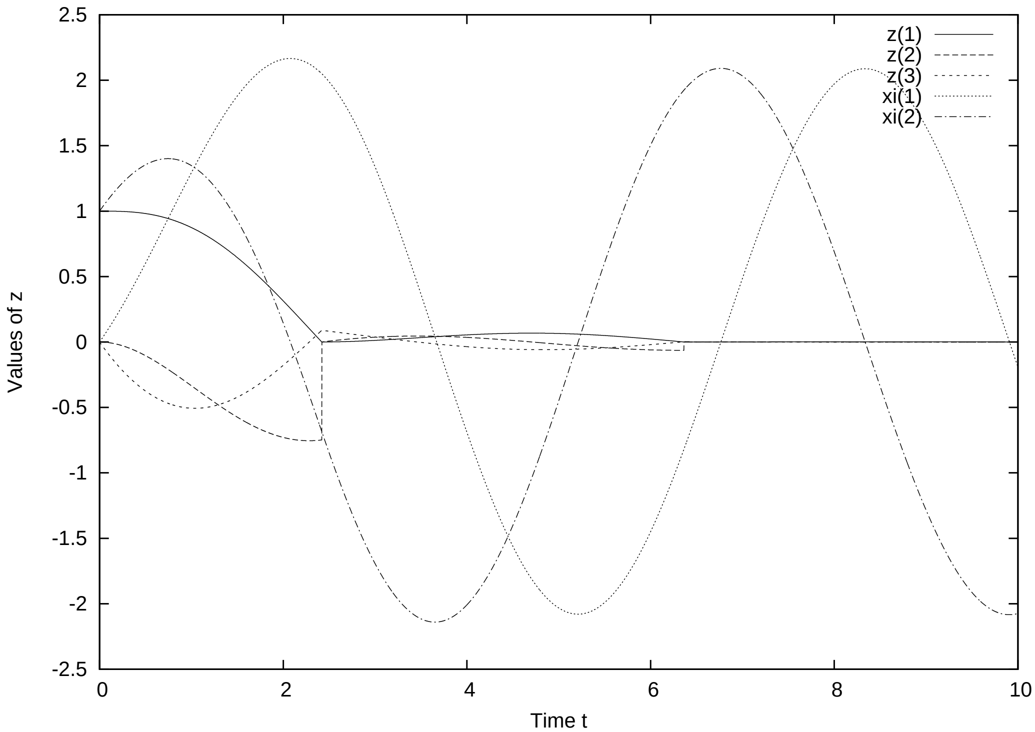


Figure 2: EMTS scheme $d_\xi = (0, 0)$, $\alpha = 1$ and $\omega = 1$. Decoupled zero-dynamics.

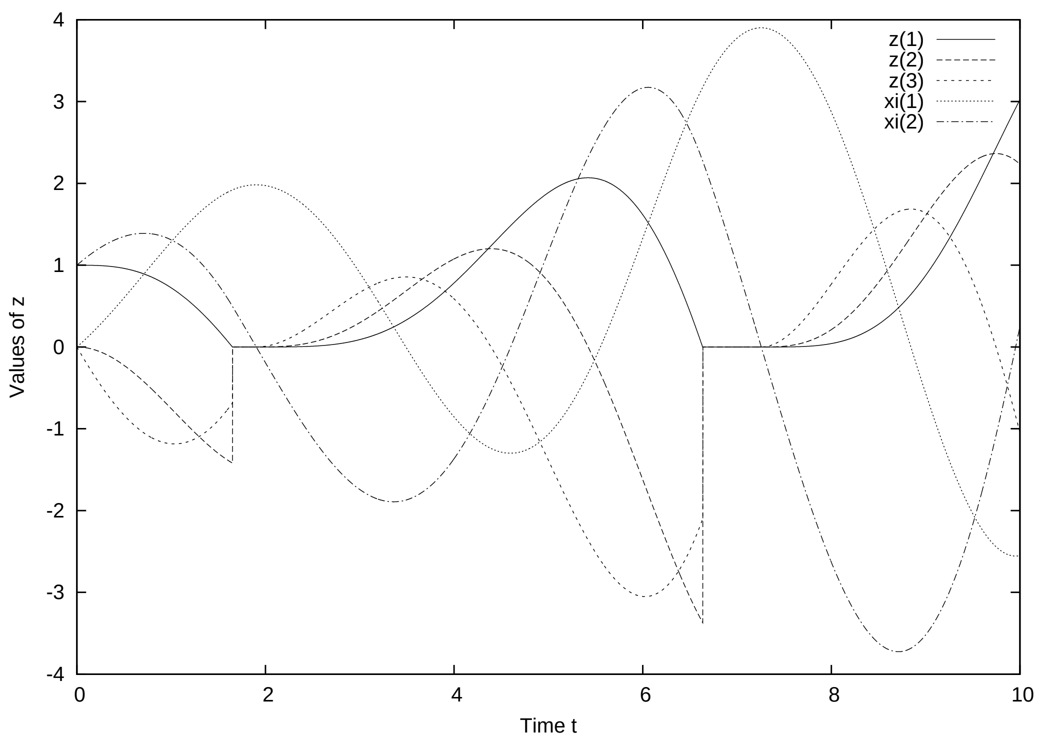


Figure 3: EMTS scheme $d_\xi = (0,1)$, $\alpha = 1$ and $\omega = 1$

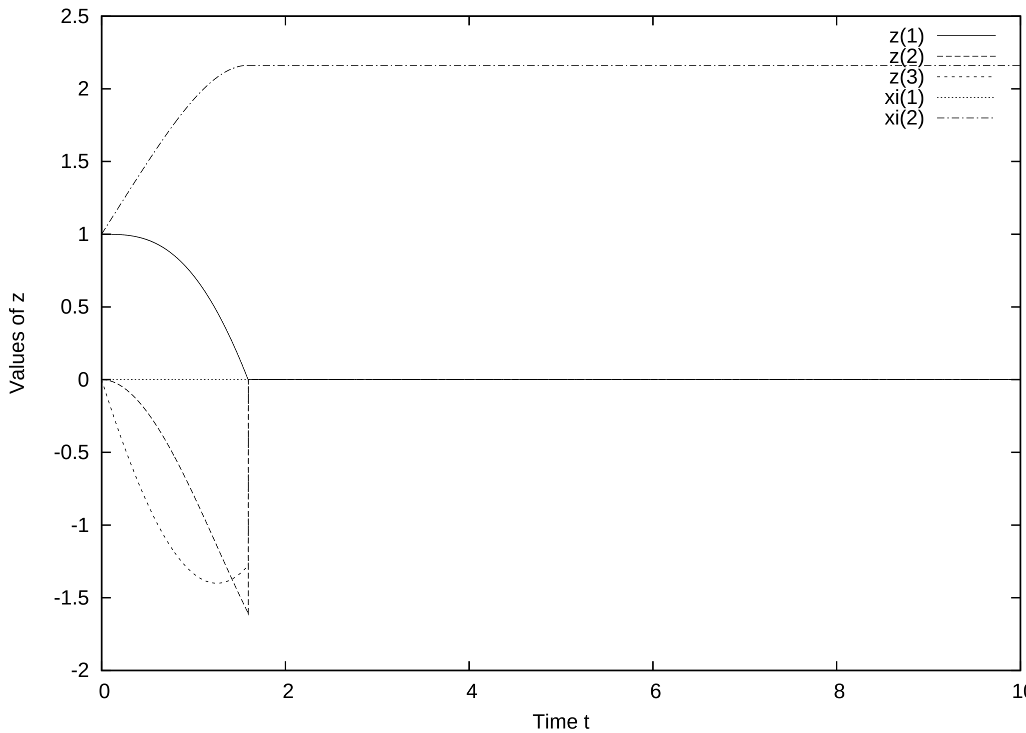


Figure 4: EMTS scheme $\alpha = 1$ and $\omega = 1$

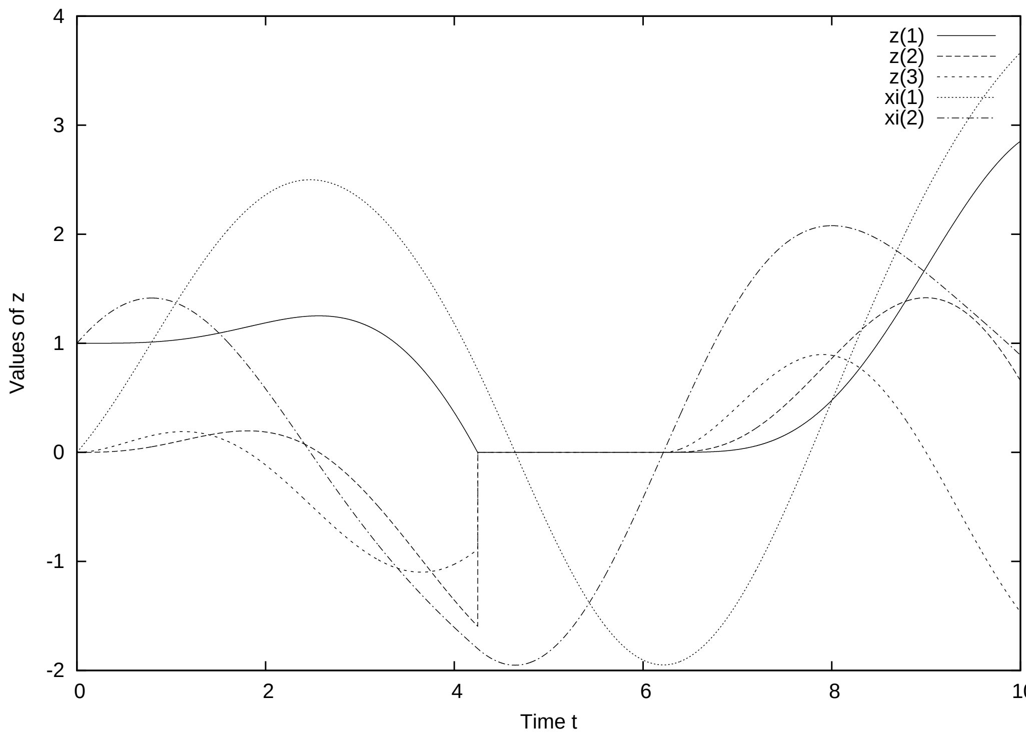


Figure 5: EMTS scheme $\alpha = 1$ and $\omega = 1$

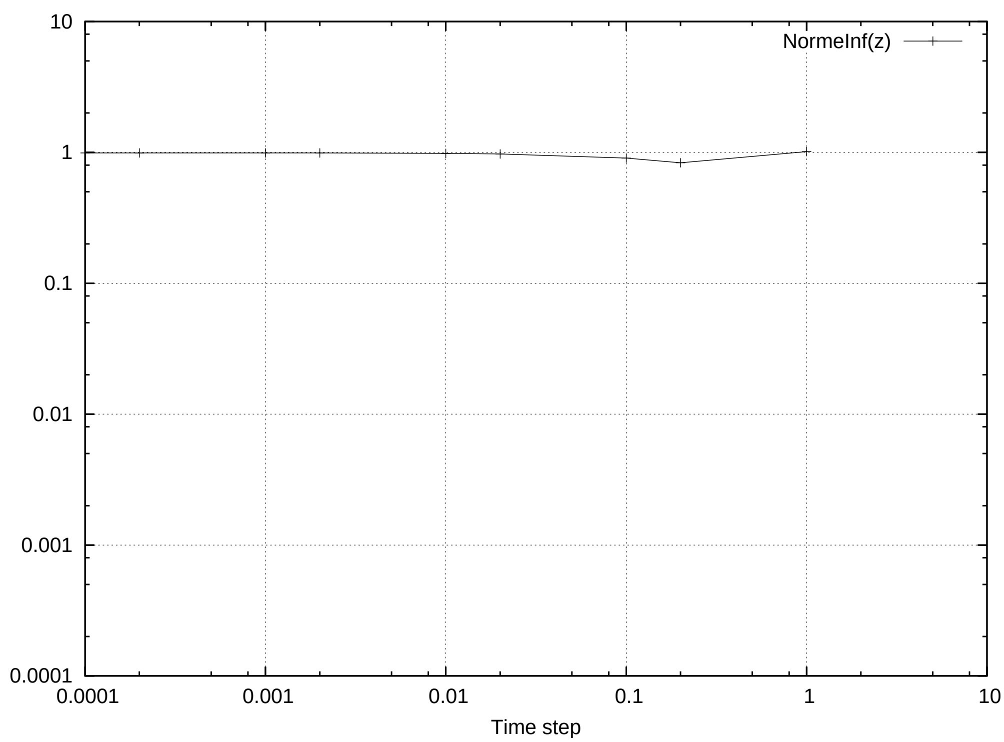


Figure 6: Empirical order of the scheme with the norm $\|.\|_\infty$

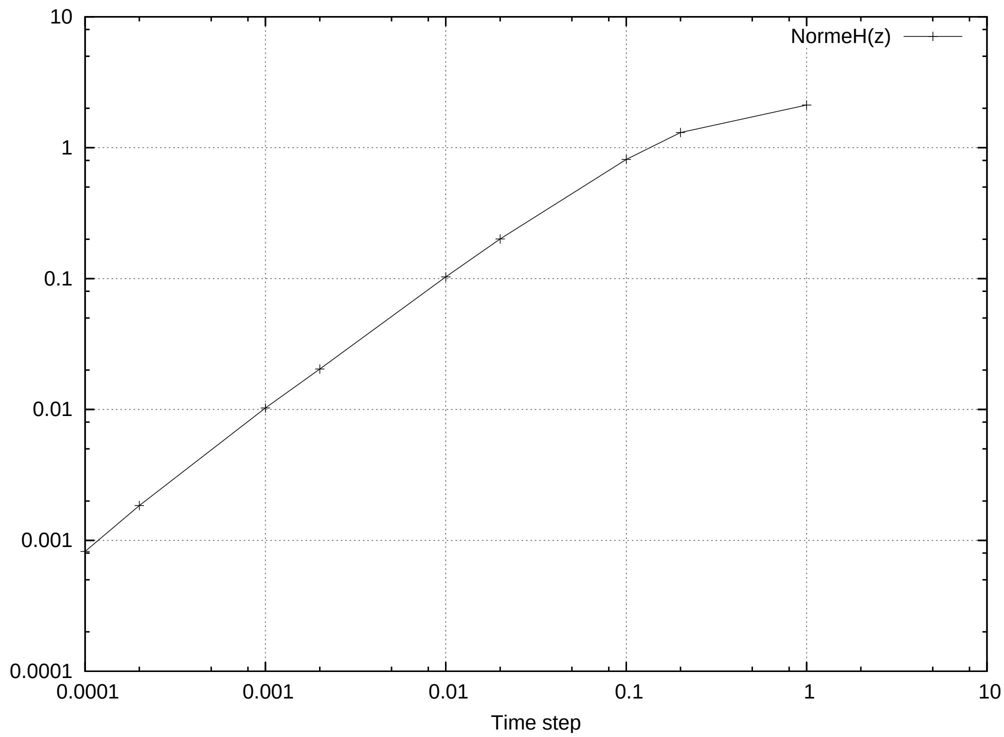


Figure 7: Empirical order of the scheme with the norm associted with the convergence in the sense of filled-in graph

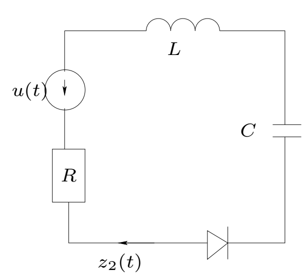


Figure 8: A simple electrical circuit with ideal diodes.

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399